INTERNATIONAL SEARCH REPORT

International Application No
PCT/FR2004/050310

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/00 C07K14/02 A61K48/00 C12Q1/68 C07K16/08
C12N15/34 A61K39/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, SEQUENCE SEARCH, BIOSIS, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENBANK 'Online! 3 June 2003 (2003-06-03) "605 bp DNA linear BCT 03-JUN-2003 Pseudomonas pseudoalcaligenes bphR2 gene for transcriptional regulator, complete cds. Pseudomonas." retrieved from NCBI Database accession no. AB088347 XP002319927 the whole document & WATANABE,T., FUJIHARA,H. AND FURUKAWA,K.: "Characterization of the Second LysR-Type Regulator in the Biphenyl-Catabolic Gene Cluster of Pseudomonas pseudoalcaligenes KF707" J. BACTERIOL., vol. 185, no. 12, June 2003 (2003-06), pages 3575-3582, XP002290937 the whole document -/--	3,4,7,8, 13,14, 25-28, 30,32,33

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 March 2005	06/04/2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Chambonnet, F

Form PCT/ISA/210 (second sheet) (January 2004)

BEST AVAILABLE COPY

INTERNATIONAL SEARCH REPORT

International Application No
PCT/FR2004/050310

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENBANK 'Online! "GSS name: pacs2-164_302.s1 Pseudomonas aeruginosa genomic clone pacs2-164_302, DNA sequence" retrieved from NCBI Database accession no. BZ561022 XP002290938 the whole document & SPENCER,D.H., RAYMOND,C.K., SMITH,E.E., SIMS,E.E., HASTINGS ,M., BURNS,J.L., KAUL,R., OLSEN,M.: "Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library" J BACTERIOL., 2002,	3,4,7,8, 19-22
P,X	WO 03/055994 A (INSERM INST NAT SANTE ET RECH ; BIO MERIEUX (FR); BEDIN FREDERIC (FR);) 10 July 2003 (2003-07-10) the whole document	3-6,11, 12, 25-28, 30, 32-49, 51-59
A	DENIS F ET AL: "LES NOUVEAUX VIRUS DES HEPATITES E, GB ET SUIVANTS. NEW HEPATITIS VIRUSES E, GB AND NEXT" TRANSFUSION CLINIQUE ET BIOLOGIQUE, ARNETTE-BLACKWELL, PARIS, FR, vol. 1, no. 3, 1996, pages 19-25, XP000612674 ISSN: 1246-7820 the whole document	1
A	BOWDEN D S ET AL: "NEW HEPATITIS VIRUSES: ARE THERE ENOUGH LETTERS IN THE ALPHABET?" MEDICAL JOURNAL OF AUSTRALIA, XX, XX, vol. 164, no. 2, 1996, pages 87-89, XP000612664 the whole document	1

BEST AVAILABLE COPY

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FR2004/050310

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

See Supplemental Sheet

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

Continuation of Box II.1

Although claim 54 relates to a diagnostic method comprising an *in vivo* surgical procedure carried out on the human body in order to take a biological sample from a patient, such as a sample of serum, plasma or blood, the search was carried out and was based on the effects ascribed to the product.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/FR2004/050310

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 03055994 A	10-07-2003	FR 2834294 A1	04-07-2003
		AU 2002364882 A1	15-07-2003
		EP 1458857 A1	22-09-2004
		WO 03055994 A1	10-07-2003
		US 2005037336 A1	17-02-2005

BEST AVAILABLE COPY

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No
PCT/FR2004/050310

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 7 C12N15/00 C07K14/02 A61K48/00 C12Q1/68 C07K16/08
C12N15/34 A61K39/12

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
CIB 7 C12N C07K A61K

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)
EPO-Internal, SEQUENCE SEARCH, BIOSIS, WPI Data

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	DATABASE GENBANK 'en ligne! 3 juin 2003 (2003-06-03) "605 bp DNA linear BCT 03-JUN-2003 Pseudomonas pseudoalcaligenes bphR2 gene for transcriptional regulator, complete cds. Pseudomonas." retrieved from NCBI Database accession no. AB088347 XP002319927 le document en entier & WATANABE,T., FUJIHARA,H. AND FURUKAWA,K.: "Characterization of the Second LysR-Type Regulator in the Biphenyl-Catabolic Gene Cluster of Pseudomonas pseudoalcaligenes KF707" J. BACTERIOL., vol. 185, no. 12, juin 2003 (2003-06), pages 3575-3582, XP002290937 le document en entier -/--	3,4,7,8, 13,14, 25-28, 30,32,33

[X] Voir la suite du cadre C pour la fin de la liste des documents

[X] Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
"E" document antérieur, mais publié à la date de dépôt international ou après cette date
"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée	Date d'expédition du présent rapport de recherche internationale
3 mars 2005	06/04/2005

Nom et adresse postale de l'administration chargée de la recherche internationale	Fonctionnaire autorisé
Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Chambonnet, F

BEST AVAILABLE COPY

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No
PCT/FR2004/050310

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	DATABASE GENBANK 'en ligne! "GSS name: pacs2-164_302.s1 Pseudomonas aeruginosa genomic clone pacs2-164_302, DNA sequence" retrieved from NCBI Database accession no. BZ561022 XP002290938 le document en entier & SPENCER,D.H., RAYMOND,C.K., SMITH,E.E., SIMS,E.E., HASTINGS ,M., BURNS,J.L., KAUL,R., OLSEN,M.: "Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library" J BACTERIOL., 2002,	3,4,7,8, 19-22
P,X	WO 03/055994 A (INSERM INST NAT SANTE ET RECH ; BIO MERIEUX (FR); BEDIN FREDERIC (FR);) 10 juillet 2003 (2003-07-10) le document en entier	3-6,11, 12, 25-28, 30, 32-49, 51-59
A	DENIS F ET AL: "LES NOUVEAUX VIRUS DES HEPATITES E, GB ET SUIVANTS. NEW HEPATITIS VIRUSES E, GB AND NEXT" TRANSFUSION CLINIQUE ET BIOLOGIQUE, ARNETTE-BLACKWELL, PARIS, FR, vol. 1, no. 3, 1996, pages 19-25, XP000612674 ISSN: 1246-7820 le document en entier	1
A	BOWDEN D S ET AL: "NEW HEPATITIS VIRUSES: ARE THERE ENOUGH LETTERS IN THE ALPHABET?" MEDICAL JOURNAL OF AUSTRALIA, XX, XX, vol. 164, no. 2, 1996, pages 87-89, XP000612664 le document en entier	1

BEST AVAILABLE COPY

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR2004/050310

Cadre II Observations – lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 2 de la première feuille)

Conformément à l'article 17.2)a), certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

1. [X] Les revendications n^{os}
se rapportent à un objet à l'égard duquel l'administration n'est pas tenue de procéder à la recherche, à savoir:

voir feuille supplémentaire SUITE DES RENSEIGNEMENTS PCT/ISA/210

2. [ ] Les revendications n^{os}
se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectuée, en particulier:

3. [ ] Les revendications n^{os}
sont des revendications dépendantes et ne sont pas rédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

Cadre III Observations – lorsqu'il y a absence d'unité de l'invention (suite du point 3 de la première feuille)

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

1. [ ] Comme toutes les taxes additionnelles ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.

2. [ ] Comme toutes les recherches portant sur les revendications qui s'y prêtaient ont pu être effectuées sans effort particulier justifiant une taxe additionnelle, l'administration n'a sollicité le paiement d'aucune taxe de cette nature.

3. [ ] Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n^{os}

4. [ ] Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n^{os}

Remarque quant à la réserve

[ ] Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant.

[ ] Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

BEST AVAILABLE COPY

SUITE DES RENSEIGNEMENTS INDIQUES SUR PCT/ISA/ 210

Suite du cadre II.1

Bien que la revendication 54 concerne une méthode de diagnostic comprenant une étape chirugicale in vivo de prélèvement d'un échantillon biologique tel que du sérum, plasma ou ou sang d'un patient appliquée au corps humain, la recherche a été effectuée et basée sur les effets imputés au produit.

BEST AVAILABLE COPY

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No
PCT/FR2004/050310

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 03055994 A	10-07-2003	FR 2834294 A1	04-07-2003
		AU 2002364882 A1	15-07-2003
		EP 1458857 A1	22-09-2004
		WO 03055994 A1	10-07-2003
		US 2005037336 A1	17-02-2005

BEST AVAILABLE COPY

SEQUENCES NUCLEIQUES ET PROTEIQUES DU VIRUS HXHV ET UTILISATIONS

L'hépatite est la plus importante des maladies transmissibles. Le mode de transmission est le plus souvent la transfusion, la transplantation d'organes et l'hémodialyse, mais l'hépatite peut également être transmise par ingestion de nourriture ou d'eau contaminée et par contact entre individus.

Les hépatites virales sont induites par divers agents viraux qui se distinguent les uns des autres par leurs génomes et leurs modes de réplication. Les hépatites virales causent des dommages au niveau du foie avec des degrés variables de sévérité. Près d'un milliard de personnes dans le monde souffrent d'hépatites virales. Il existe des risques graves dans les formes chroniques des hépatites qui peuvent évoluer en cirrhose ou hépatocarcinome. Les hépatites virales peuvent être diagnostiquées par mise en évidence de symptômes bien définis, tels que la jaunisse, des taux élevés de transaminases (aspartate transaminase ou AST, alanine transaminase ou ALT, lactate déshydrogénase ou LDH), et des lésions hépatiques. Mais malgré la connaissance de différents virus des hépatites A, B, C, D, E, G et TTV, 5% de toutes les hépatites et 40% des hépatites fulminantes demeurent inexpliquées, d'où l'hypothèse de l'existence de virus inconnus des hépatites. Ces hépatites sans étiologie connue sont aussi bien post-transfusionnelles que sporadiques, chroniques ou fulminantes. Elles sont communément appelées hépatites X.

Les virus des hépatites G (GBV-A, GBV-B, GBV-C) et TTV récemment identifiés ne semblent pas être pathogènes chez l'homme et ne peuvent donc pas expliquer les cas d'hépatites sans étiologie connue ou hépatites X.

A partir d'un cas d'hépatite grave sans étiologie connue, d'un patient chez lequel un traitement à l'interféron a permis de normaliser les transaminases, un nouveau virus dénommé HXHV, associé aux hépatites X, a été décrit. Le génome du virus HXHV est un génome à ADN au

moins partiellement simple brin qui comprend une ou plusieurs trames de lecture codant pour une ou des protéine(s) ou polyprotéine(s), le génome comprend une séquence nucléotidique susceptible de s'hybrider à la séquence nucléotidique XH ou à la séquence nucléotidique complémentaire de la séquence XH. La séquence XH est représentée dans l'identificateur de séquences de la présente demande en SEQ ID NO :1. La séquence XH est riche en GC (62%) et présente quatre trames de lecture ouvertes (ORF1, ORF2, ORF3, ORF4). Cette séquence isolée a été caractérisée et aucune homologie de séquences avec l'ADN génomique humain et avec les séquences présentes dans les bases de données n'a été retrouvée. Toutes les informations concernant le virus HXHV sont contenues dans la demande de brevet PCT/FR02/04578 déposée aux noms des demanderesses.

Les présents inventeurs ont maintenant isolé et caractérisé une nouvelle séquence nucléotidique du virus HXHV. Cette séquence, dénommée XH1 est riche en GC (61,2%), ce qui est comparable avec la teneur en GC de la séquence XH isolée précédemment. La séquence XH1 est référencée dans l'identificateur de séquences en SEQ ID NO : 4. La séquence XH1 ne présente aucune homologie ou identité significative avec toutes les séquences disponibles dans les bases de données. Elle présente 5 trames de lecture ouvertes. Les séquences ADN correspondant auxdites trames de lecture ouvertes sont respectivement identifiées en SEQ ID NOs 5 à 9 dans l'identificateur de séquences. Comme il est de nature courante dans le domaine de la virologie, les présents inventeurs ont généré le brin ADN complémentaire de la séquence XH1 et ont également recherché s'il existait de potentielles trames de lecture ouvertes sur le brin ADN complémentaire. Ils ont identifiés 8 trames de lecture ouvertes qui sont respectivement représentées en SEQ ID NOs 10 à 17. Les séquences polypeptidiques correspondant

auxdites trames de lecture sont respectivement identifiées en SEQ ID NOs : 18 à 30 dans l'identificateur de séquences. Les séquences précitées et leurs fragments sont utilisés pour la détection du virus HXHV.

Ainsi, la présente invention concerne :

- une séquence d'acide nucléique susceptible d'être obtenue à partir du génome du virus HXHV, ladite séquence d'acide nucléique comprenant ou consistant en SEQ ID NO : 4.

- un fragment nucléotidique d'ADN isolé comprenant ou consistant en une séquence nucléotidique ADN ou ARN d'au moins 12 nucléotides contigus, de préférence d'au moins 15 ou d'au moins 18 nucléotides contigus, et avantageusement d'au moins 20, 21, 22, 23, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51 ou 54 nucléotides contigus, de la séquence nucléotidique ADN SEQ ID NO : 4 ou de la séquence ADN complémentaire de SEQ ID NO : 4 ; ou d'une séquence nucléotidique qui présente, sur au moins 12 nucléotides contigus, de préférence sur au moins 15 ou au moins 18 nucléotides contigus, et avantageusement sur au moins 20, 21, 22, 23, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51 ou 54 nucléotides contigus, au moins 90%, de préférence au moins 92%, 95% ou au moins 98%, 99% d'homologie ou d'identité par rapport à la séquence représentée en SEQ ID NO : 4 ou par rapport à séquence ADN complémentaire de SEQ ID NO : 4 ; à l'exclusion des fragments qui consistent en une des séquences nucléotidiques suivantes : TAGTCGAGACTCAACCATCGC, CCCGCCCCGCTGATGAAAAG et des séquences nucléotidiques complémentaires desdites séquences ; ou à la condition que sur 20 nucléotides ou 21 nucléotides contigus ledit fragment nucléotidique d'ADN ne présente pas 100% d'homologie ou d'identité avec un fragment nucléotidique de la séquence référencée SEQ ID NO : 1 ou de la séquence complémentaire de SEQ ID NO : 1. Ledit fragment est en particulier choisi parmi les fragments dont lesdits nucléotides contigus appartiennent

à l'un des segments suivants : un segment dont la séquence commence au nucléotide 2 et se termine au nucléotide 286 de SEQ ID NO :4, un segment dont la séquence commence au nucléotide 4 et se termine au nucléotide 144 de SEQ ID NO :4, un segment dont la séquence commence au nucléotide 180 et se termine au nucléotide 1004 de SEQ ID NO :4, un segment dont la séquence commence au nucléotide 614 et se termine au nucléotide 820 de SEQ ID NO :4, un segment dont la séquence commence au nucléotide 1228 et se termine au nucléotide 1314 de SEQ ID NO :4 ou les fragments complémentaires ; un segment dont la séquence commence au nucléotide 1283 et se termine au nucléotide 1197 de la séquence complémentaire de SEQ ID NO :4, un segment dont la séquence commence au nucléotide 1264 et se termine au nucléotide 1067 de la séquence complémentaire de SEQ ID NO : 4, un segment dont la séquence commence au nucléotide 1209 et se termine au nucléotide 1099 de la séquence complémentaire de SEQ ID NO : 4, un segment dont la séquence commence au nucléotide 819 et se termine au nucléotide 736 de la séquence complémentaire de SEQ ID NO : 4, un segment dont la séquence commence au nucléotide 800 et se termine au nucléotide 6 de la séquence complémentaire de SEQ ID NO : 4, un segment dont la séquence commence au nucléotide 784 et se termine au nucléotide 629 de la séquence complémentaire de SEQ ID NO : 4, un segment dont la séquence commence au nucléotide 610 et se termine au nucléotide 410 de la séquence complémentaire de SEQ ID NO : 4, un segment dont la séquence commence au nucléotide 391 et se termine au nucléotide 221 de la séquence complémentaire de SEQ ID NO : 4 ou les fragments complémentaires ; et de préférence un fragment comprenant ou consistant en l'une quelconque des séquences SEQ ID NOs : 5 à 17 ou en l'une quelconque des séquences ADN complémentaires de SEQ ID NO : 5 à 17 (le segment dont la séquence commence au nucléotide 180 et

se termine au nucléotide 1004 de SEQ ID NO :4 code pour une protéine des transposase/intégrases);

- le produit de transcription la séquence comprenant ou consistant en SEQ ID NO : 4 ou le produit de transcription d'un fragment tel que défini ci-dessus, ou le produit de transcription de la séquence comprenant ou consistant en la séquence complémentaire de SEQ ID NO :4 ;

- une molécule d'ADN qui comprend ou consiste en une séquence nucléotidique ADN représentée en SEQ ID NO : 4 ou en ce qu'elle comprend au moins un fragment nucléotidique ADN tel que défini ci-dessus ou leurs séquences complémentaires ;

- une molécule d'ARN qui comprend ou consiste en une séquence nucléotidique ARN qui est le produit de transcription d'une séquence nucléotidique ADN représentée en SEQ ID NO : 4 ou de la séquence complémentaire de SEQ ID NO : 4 ou qui est le produit de transcription d'au moins un fragment tel que défini ci-dessus ou leurs séquences complémentaires.

L'homologie et identité ci-dessus recouvre et les équivalents fonctionnels de la séquence SEQ ID NO : 4, c'est à dire les séquences ADN dans lesquelles au moins un codon peut être remplacé par un autre codon tout en codant pour un acide aminé identique. On parle de dégénérescence du code génétique. Ainsi, les codes de l'argininine, de la sérine et de la leucine présentent une dégénérescence d'ordre 6 (c'est à dire qu'il y a six codons différents pour chacune d'elle), tandis que les codes d'autres acides aminés, tels que l'acide glutamique, la glutamine, la tyrosine, l'histidine et quelques autres présentent une dégénérescence d'ordre 2. De tous les acides aminés seuls le tryptophane et la méthionine ont une dégénérescence d'ordre 1. Il est donc clair que pour l'expression d'un polypeptide dont la séquence est représentée en SEQ ID NOs : 18 à 30, on peut utiliser des séquences d'acides nucléiques variantes et fonctionnelles dont les

compositions en codons sont différentes de la séquence d'acide nucléique représentée en SEQ ID NO : 4 ou de sa séquence complémentaire.

L'homologie ou identité définie ci-dessus vise également les variants du virus HXHV et les séquences mutantes du virus HXHV, et en particulier celles issues de la variabilité naturelle. En effet, il est bien connu des spécialistes que les virus ont des taux relativement élevés de mutations spontanées ou induites.

L'invention concerne également :

- un polypeptide comprenant une séquence polypeptidique codée par une séquence ou par un fragment tel(le) que défini(e) ci-dessus ou par leurs équivalents fonctionnels ou par une séquence nucléotidique qui présente au moins 90% d'homologie ou d'identité, de préférence au moins 92% ou 95% d'homologie ou d'identité et avantageusement au moins 98% ou 99% d'homologie ou d'identité par rapport à la séquence représentée en SEQ ID NO : 4 ou par rapport à la séquence complémentaire de SEQ ID NO : 4, à la condition que les séquences TAGTCGAGACTCAACCATCGC, CCCGCCCCGCTGATGAAAAG, les séquences nucléotidiques complémentaires desdites séquences soient exclues ; ou à la condition que sur 20 nucléotides ou 21 nucléotides contigus le fragment nucléotidique d'ADN ne présente pas 100% d'homologie ou d'identité avec un fragment nucléotidique de la séquence référencée SEQ ID NO : 1 ou de la séquence complémentaire de SEQ ID NO : 1;

- un polypeptide dont la séquence polypeptidique comprend ou consiste en l'une quelconque des séquences SEQ ID Nos 18 à 30 ou en une séquence polypetidique fonctionnellement équivalente auxdites séquences ;

- un fragment polypeptidique, caractérisé en ce qu'il comprend ou consiste en une séquence peptidique d'au moins 4 acides aminés contigus, de préférence d'au moins 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 ou 18 acides aminés de l'une quelconque des séquences peptidiques

représentées en SEQ ID NO : 18 à 30 ou d'une séquence peptidique fonctionnellement équivalente auxdites séquences SEQ ID NO : 18 à 30 ; étant entendu que par séquence peptidique fonctionnellement équivalente on entend une séquence peptidique qui est reconnue par des anticorps dirigés contre le virus HXHV ;

- un fragment polypeptidique qui comprend ou qui consiste en une séquence peptidique représentée en l'une quelconque des SEQ ID NOs : 18 à 30 ou une séquence peptidique fonctionnellement équivalente à l'une quelconque des SEQ ID NOs : 18 à 30 ; étant entendu que par séquence peptidique fonctionnellement équivalente on entend une séquence peptidique qui est reconnue par des anticorps dirigés contre le virus HXHV.

- un épitope caractérisé en ce qu'il comprend ou consiste en une séquence peptidique d'au moins 6, 8, 9, 10, 12, 15 ou 18 acides aminés et au plus de 10, 12, 15 ou 18 acides aminés, en particulier en ce que sa séquence consiste en une séquence peptidique de 6 à 10 acides aminés, de 6 à 12 acides aminés, de 6 à 15 acides aminés, de 6 à 18 acides aminés, de 8 à 10 acides aminés, de 8 à 12 acides aminés, de 8 à 15 acides aminés, de 8 à 18 acides aminés et de 15 à 18 acides aminés de l'une quelconque des séquences représentées en SEQ ID NO : 18 à 30 ou d'une séquence polypeptidique fonctionnellement équivalente auxdites séquences SEQ ID NO : 18 à 30 ; étant entendu que ledit épitope est reconnu par des anticorps dirigés contre le virus HXHV.

Par "polypeptide", on désigne un peptide, à l'état isolé, présentant un enchaînement d'un nombre variable d'acides aminés, tel qu'un oligopeptide, une protéine, une protéine de fusion, un peptide de fusion, un peptide de synthèse. Un polypeptide peut être obtenu par différentes techniques bien connues de l'homme du métier, et notamment par synthèse chimique ou par des techniques de recombinaison génétique. Les polypeptides selon

l'invention peuvent être obtenus par des méthodes de synthèse classique, par exemple avec un synthétiseur automatique de peptides, ou par les techniques de génie génétique comprenant l'insertion d'une séquence d'ADN codant pour ledit polypeptide dans un vecteur d'expression tel qu'un plasmide ou un virus, et la transformation de cellules avec ce vecteur d'expression et culture de ces cellules.

Par séquence peptidique fonctionnellement équivalente à une séquence peptidique de référence, on entend une séquence d'acides aminés modifiée par insertion et/ou délétion et/ou substitution et/ou allongement et/ou raccourcissement et/ou modification chimique d'un ou plusieurs acides aminés, pour autant que ces modifications préservent substantiellement voire développent les propriétés immunoréactives de ladite séquence peptidique de référence.

Ainsi, on entend par séquences fonctionnellement équivalentes des séquences qui conservent les propriétés immunoréactives de SEQ ID Nos 18 à 30 ou de leurs fragments, notamment les séquences dans lesquelles un ou plusieurs acide(s) aminé(s) est ou sont substitué(s) par un ou plusieurs autres acides aminés ; les séquences dans lesquelles un ou plusieurs acide aminé de la série L est remplacé par un acide aminé de la série D, et vice-versa ; les séquences dans lesquelles on a introduit une modification des chaînes latérales des acides aminés, telle qu'une acétylation des fonctions amines, une carboxylation des fonctions thiols, une estérification des fonctions carboxyliques ; une modification des liaisons peptidiques telles que par exemple des liaisons carba, rétro, inverso, retro-inverso, réduites et méthylène-oxy.

Par exemple un ou plusieurs acide(s) aminé(s) dans les séquences des polypeptides de l'invention peuvent être substitué(s) par un ou plusieurs autre(s) acide(s) aminé(s) de polarité similaire qui agissent comme des

équivalents fonctionnels. Des substitutions pour un acide aminé dans des séquences polypeptiques d'intérêt peuvent être déterminées à partir d'autres membres de la classe auquel l'acide aminé appartient. Par exemple, les acides aminés non polaires (hydrophobes) comprennent l'alanine, la leucine, l'isoleucine, la valine, la proline, la phénylalanine, la tryptophane, la méthionine. Les acides aminés neutres polaires comprennent la glycine, la sérine, la thréonine, la cystéine, la tyrosine, l'asparagine, la glutamine. Les acides aminés chargés positivement (basiques) comprennent l'arginine, la lysine et l'histidine. Les acides aminés chargés négativement (acides) comprennent l'acide aspartique et l'acide glutamique. D'autres substitutions pour un acide aminé dans des séquences polypeptidiques d'intérêt peuvent être déterminées à partir des informations contenues dans l'article de Kramer A. et al. (Molecular Immunology, Vol. 32, N°7, pp. 459-465 (1995)). Ces auteurs ont constitué des banques dans lesquelles pour réduire le problème de l'explosion combinatoire du nombre de molécules, ils ont utilisés des groupes d'acides aminés constitués d'acides aminés ayant des propriétés physico-chimiques similaires et ce sont les acides aminés regroupés dans chacun de ces six groupes, listés ci-dessous, qui sont considérés principalement comme équivalents dans la présente invention.

Groupe 1 : alanine, proline, glycine.
Groupe 2 : acide aspartique, acide glutamique.
Groupe 3 : histidine, lysine, arginine.
Groupe 4 : asparagine, glutamine, sérine, thréonine.
Groupe 5 : phénylalanine, tyrosine, tryptophane.
Groupe 6 : isoleucine, leucine, valine, méthionine.

L'équivalence d'une séquence peptidique par rapport à une séquence peptidique de référence peut être définie par son identité ou son homologie, exprimée en pourcentage, avec ladite séquence de référence. Ce pourcentage est

déterminé, pour une suite d'un nombre donné d'acides aminés contigus, par alignement des deux séquences, déplacement de l'une par rapport à l'autre, et comparaison des acides aminés dans les deux séquences. Le pourcentage d'identité est déterminé à partir du nombre d'acides aminés qui sont identiques à des acides aminés de la séquence de référence, dans la même position. Le pourcentage d'homologie est déterminé à partir du nombre d'acides aminés qui sont équivalents à des acides aminés de la séquence de référence, dans la même position.

L'invention concerne aussi une cassette d'expression fonctionnelle dans une cellule issue d'un organisme procaryote ou eucaryote permettant l'expression d'une séquence d'acide nucléique ou d'un fragment d'ADN ou d'une molécule d'ADN tels que décrits ci-dessus, placé sous le contrôle des éléments nécessaires à son expression. La cassette d'expression est caractérisée en ce qu'elle est fonctionnelle dans une cellule issue d'un organisme procaryote, en particulier *E. coli* ou d'un organisme eucaryote, en particulier les cellules provenant d'animaux tels que les mammifères, les reptiles, les insectes, en particulier les cellules COS, CHO, BHK, PK 15, RK 13 ; les lignées cellulaires humaines de l'ostéosarcorme (cellules 143 B), les lignées cellulaires humaines HeLa et les lignées cellulaires humaines de l'hépatome (du type HepG2) ; les lignées cellulaires d'insecte (par exemple de Spodoptera frugiperda) ; ou eucaryote inférieur, en particulier les cellules de levures, telles que *Saccharomyces, Schizosaccharomyces, Kluveromyces, Hanseluna, Yarowia, Schwaniomyces, Zygosaccharomyces* et *Pichia*, et de préférence choisi parmi les cellules de *Saccharomyces cerevisiae, Saccharomyces carlsbergensis, Schizosaccharomyces pombe, Kluveromyces lactis* et de *Pichia pastoris*.

L'invention concerne encore un vecteur comprenant ladite cassette d'expression ; une cellule issue d'un

organisme procaryote, eucaryote ou eucaryote inférieur , de préférence un organisme eucaryote ou eucaryote inférieur tel que défini ci-dessus ou un vecteur tel que défini ci-dessus ; et le polypeptide susceptible d'être produit par la cassette d'expression, le vecteur ou la cellule.

L'invention a pour objet un procédé pour préparer un polypeptide ou un fragment polypeptidique tel que défini ci-dessus qui consiste à cultiver une cellule hôte répondant aux définitions précédentes dans un milieu de culture approprié, ladite cellule hôte étant transformée avec un vecteur d'expression qui contient une séquence d'acide nucléique ADN telle que définie précédemment ou un fragment nucléotidique ADN tel que défini précédemment ou une molécule d'ADN telle que définie précédemment et, à purifier ledit polypeptide produit jusqu'à un degré de pureté requis.

L'invention a aussi pour objet un polypeptide immunogène, ledit polypeptide comprenant ou consistant en une séquence polypeptidique ou peptidique telle que définie précédemment. Un tel polypeptide immunogène est utilisé pour la production d'anticorps monoclonaux ou polyclonaux ou de fragments desdits anticorps et l'invention englobe les anticorps monoclonaux ou polyclonaux ou leurs fragments, étant obtenus par immunisation d'un animal mammifère (lapin, rat, souris) avec un tel peptide immunogène.

La production d'anticorps monoclonaux ou polyclonaux est bien connue de l'homme du métier. On peut citer à titre de référence Köhler G. et Milstein C. (1975) : Continuous culture of fused cells secreting antibody of predefined specificity, Nature 256 :495-497 et Galfre G. et al. (1977) Nature, 266 : 522-550 pour la production d'anticorps monoclonaux et Roda A., Bolelli G.F. : Production of high-titer antibody to bile acids, Journal of Steroid Biochemistry, Vol. 13, pp 449-454 (1980) pour

la production d'anticorps polyclonaux. Des anticorps peuvent également être produits par immunisation de souris, de rat ou de lapins avec les particules virales de HXHV. Pour la production d'anticorps polyclonaux et monoclonaux, l'immunogène peut être couplé à de l'albumine sérique (peptide SA) ou à l'hémocyanine de Lymphet Keyhole (peptide KLH) comme support pour l'immunisation. Les anticorps sont ensuite criblés pour leur spécificité en utilisant les techniques habituelles, telles que des tests ELISA ou de Western Blot. Pour la production d'anticorps monoclonaux les animaux sont soumis à une injection d'immunogène en utilisant de l'adjuvant complet de Freund. Les sérums et les surnageants de culture d'hybridome issus des animaux immunisés sont analysés pour leur spécificité et leur sélectivité en utilisant des techniques classiques, telles que par exemple des tests ELISA ou de Western Blot. Les hybridomes produisant les anticorps les plus spécifiques et les plus sensibles sont sélectionnés. Des anticorps monoclonaux peuvent également être produits *in vitro* par culture cellulaire des hybridomes produits ou par récupération de liquide d'ascite, après injection intrapéritonéale des hybridomes chez la souris. Quel que soit le mode de production, en surnageant ou en ascite, les anticorps sont ensuite purifiés. Les méthodes de purification utilisées sont essentiellement la filtration sur gel échangeur d'ions et la chromatographie d'exclusion ou la chromatographie d'affinité (protéine A ou G). Un nombre suffisant d'anticorps sont criblés dans des tests fonctionnels pour identifier les anticorps les plus performants. La production *in vitro* d'anticorps, de fragments d'anticorps ou de dérivés d'anticorps, tels que des anticorps chimères produits par génie génétique est bien connue de l'homme du métier. Il est avantageux d'utiliser des anticorps humanisés. Les formes " humanisées " d'anticorps non humains, par exemple murins, sont des anticorps chimères qui comprennent une

séquence minimale dérivée d'une immunoglobuline non humaine. Pour la plupart, les anticorps humanisés sont des immunoglobulines humaines (anticorps récepteur) dans lesquelles des résidus d'une région hypervariable du récepteur sont remplacés par des résidus d'une région hypervariable d'une espèce donneur (anticorps donneur) non humaine, telle que souris, rat, lapin ou primate non humain, ayant la spécificité, l'affinité et la capacité souhaitées. Dans certains cas, les résidus (FR) de la région Fv de l'immunoglobuline humaine sont remplacés par des résidus correspondants non humains. De plus, les anticorps humanisés peuvent comprendre des résidus qui ne sont pas trouvés dans l'anticorps receveur ou dans l'anticorps donneur. Ces modifications sont effectuées pour améliorer les performances de l'anticorps. En général, l'anticorps humanisé comprendra au moins et de préférence deux domaines variables, dans lesquels tout ou à peu près tout des boucles hypervariables correspondent à une immunoglobuline non humaine et tout ou à peu près tour des régions FR seront celles d'une immunoglobuline humaine. Les anticorps humanisés facultativement pourront aussi comprendre au moins une partie d'une région constante (Fc) d'une immunoglobuline, telle qu'une immunoglobuline humaine (Jones et al., Nature 321 : 522-525 (1986) ; Reichmann et al., Nature 332 : 323-329 (1988) ; et Presta et al., Curr. Op. Struct. Biol. 2 : 593-596 (1992).

Plus particulièrement, par fragment d'anticorps on entend les fragments F(ab)2, Fab, Fab', sFv (Blazar et al., 1997, Journal of Immunology 159 : 5821-5833 et Bird et al., 1988, Science 242 : 423-426) d'un anticorps natif et par dérivé on entend, entre autres, un dérivé chimérique d'un anticorps natif (voir par exemple Arakawa et al., 1996, J. Biochem 120 : 657-662 et Chaudray et al., 1989, Nature 339 : 394-397). Ces fragments d'anticorps et

dérivés d'anticorps conservent la capacité de se lier sélectivement à l'antigène cible.

L'anticorps monoclonal ou polyclonal ainsi obtenu ou son fragment est incorporé dans une composition diagnostique qui est utilisée dans un procédé pour détecter au moins un polypeptide ou un fragment peptidique tel que défini précédemment dans un échantillon biologique, selon lequel on met en contact l'échantillon biologique avec la composition dans des conditions prédéterminées qui permettent la formation de complexes anticorps/antigène et on détecte la formation desdits complexes.

L'invention a également pour objet une composition diagnostique qui comprend un polypeptide ou un fragment peptidique tel que défini précédemment et un procédé pour détecter des anticorps dirigés contre le virus HXHV ou au moins contre un polypeptide ou un fragment peptidique de l'invention, selon lequel on met en contact un échantillon biologique suspecté être ou pouvoir avoir été infecté par le virus HXHV avec la composition diagnostique dans des conditions prédéterminées qui permettent la formation de complexes anticorps/antigène et on détecte la formation desdits complexes. En effet, il est connu que lors d'une infection par un agent viral, l'hôte développe des anticorps dirigés contre cet agent viral (réponse humorale).

La présente invention a aussi pour objet le matériel biologique pour la préparation d'une composition pharmaceutique destinée au traitement des êtres humains ou des animaux infectés par au moins le virus HXHV et des compositions immunogènes ou vaccinales qui peuvent être utilisées pour produire des vaccins thérapeutiques contre une infection par le virus HXHV et des vaccins prophylactiques pour prévenir une potentielle infection par le virus HXHV, lesdites préparations immunogènes comprenant au moins un polypeptide ou un fragment

peptidique naturel, recombinant, ou de synthèse de l'invention associé à un véhicule et/ou un adjuvant et/ou un diluant et/ou un excipient pharmaceutiquement acceptable.

La présente invention a aussi pour objet l'utilisation d'au moins un anticorps monoclonal ou polyclonal ou d'au moins un fragment desdits anticorps de l'invention, spécifique d'au moins un polypeptide ou un fragment peptidique tel que défini ci-dessus pour la préparation d'une composition pharmaceutique qui administrée à un patient infecté par le virus HXHV a la capacité de réduire voire d'inhiber la prolifération et/ou la réplication du virus. Ces anticorps ou leurs fragments sont appelés anticorps neutralisants.

Par échantillon biologique, on entend par exemple le sang, le sérum, le plasma, les prélèvements tissulaires, tels que les extraits de biopsie du foie.

Les vaccins préparés sont injectables, c'est-à-dire en solution liquide ou en suspension. En option, la préparation peut aussi être émulsifiée. La molécule antigénique peut être mélangée avec des excipients qui sont pharmaceutiquement acceptables et compatibles avec l'ingrédient actif. Des exemples d'excipients favorables sont l'eau, une solution saline, le dextrose, le glycérol, l'éthanol ou des équivalents et leurs combinaisons. Si désiré, le vaccin peut contenir des quantités mineures de substances auxiliaires comme des agents mouillants ou émulsifiants, des agents qui tamponnent le pH ou des adjuvants comme l'hydroxide d'aluminium, le dipeptide muramyl ou leurs variations. Dans le cas des peptides, leur couplage à une plus grosse molécule (KLH, toxine tétanique) augmente quelquefois l'immunogénicité. Les vaccins sont administrés conventionnellement par injection par exemple intramusculaire. Des formulations additionnelles favorables avec d'autres modes

d'administration incluent des suppositoires et quelquefois des formulations orales.

Par " véhicule pharmaceutiquement acceptable " on entend les supports et véhicules administrables à l'être humain ou à un animal, tels que décrits par exemple dans Remington's Pharmaceutical Sciences 16th ed., Mack Publishing Co. Le véhicule pharmaceutiquement acceptable est de préférence isotonique, hypotonique ou présente une faible hypertonicité et a une force ionique relativement basse. Les définitions des excipients et adjuvants pharmaceutiquement acceptables sont également données dans Remington's Pharmaceutical Sciences précité.

L'invention a encore pour objet :

- une sonde, caractérisée en ce qu'elle est susceptible de s'hybrider dans des conditions de stringence déterminées à une séquence d'acide nucléique ou à un fragment nucléotidique d'ADN ou d'ARN ou à une molécule d'ADN ou d'ARN de l'invention ; de préférence, une sonde de l'invention comprend au moins 12 nucléotides, de préférence au moins 17, 18, 19, 20, 21, 22, 23, 24 ou 25 nucléotides et l'hybridation est réalisée dans des conditions de stringence correspondant à une combinaison de la température et de la concentration saline choisie approximativement entre 12 à 20°C sous le Tm (« melting temperature ») du complexe sonde / séquence nucléotidique à détecter ;

- une amorce, caractérisée en ce qu'elle est susceptible de s'hybrider dans des conditions de stringence déterminées à une séquence d'acide nucléique ou à un fragment nucléotidique d'ADN ou d'ARN ou à une molécule d'ADN ou d'ARN de l'invention ; de préférence, une amorce de l'invention comprend au moins 12 nucléotides, de préférence au moins 17, 18, 19, 20, 21, 22, 23, 24 ou 25 nucléotides et l'hybridation est réalisée dans des conditions de stringence correspondant à une combinaison de la température et de la concentration

saline choisie approximativement entre 12 à 20°C sous le Tm (« melting temperature ») du complexe amorce / séquence nucléotidique à amplifier et/ou détecter. Les amorces représentées en SEQ ID Nos 32 à 37 sont nouvelles et comme décrit dans la partie expérimentale des couples d'amorces sont utilisés pour l'amplification des acides nucléiques du virus HXHV, lesdits couples d'amorces étant choisis préférentiellement parmi les couples suivants : SEQ ID NO : 31 / SEQ ID NO : 32, SEQ ID NO : 31 / SEQ ID NO : 33, SEQ ID NO : 34 / SEQ ID NO : 35, , SEQ ID NO : 36 / SEQ ID NO : 37 ;

- un anticorps anti-acide nucléique, caractérisé en ce qu'il est susceptible de se lier à une séquence d'acide nucléique ou à un fragment nucléotidique d'ADN ou d'ARN ou à une molécule d'ADN ou d'ARN ;

- une composition diagnostique, caractérisée en ce qu'elle comprend au moins une sonde ou une amorce ou un anticorps anti-acide nucléique tel que défini ci-dessus ;

- un procédé de détection d'ADN et/ou d'ARN viral, selon lequel on prélève un échantillon biologique d'un patient suspecté être ou pouvoir avoir été infecté par le virus HXHV, on traite si nécessaire ledit échantillon pour en extraire l'ADN et/ou l'ARN, on met en contact ledit échantillon avec au moins une sonde ou une amorce de l'invention, dans des conditions de stringence déterminées et on détecte la présence d'ADN et/ou d'ARN viral dans l'échantillon soit par mise en évidence d'une hybridation dudit ADN et/ou ARN viral avec au moins une sonde, soit par amplification dudit ADN et/ou ARN (par exemple comme décrit dans la partie expérimentale de l'invention) ; et

- un procédé de détection d'ADN et/ou d'ARN viral, selon lequel on prélève un échantillon de sérum ou de plasma d'un patient, on traite si nécessaire ledit échantillon pour en extraire l'ADN et/ou l'ARN, on met en contact ledit échantillon avec au moins un anticorps anti-acide nucléique, ledit anticorps étant éventuellement

marqué par tout marqueur approprié et on met en évidence la formation d'un complexe acide nucléique/anticorps.

La production de polynucléotides, sondes ou amorces fait partie des connaissances générales de l'homme du métier. On peut notamment citer l'utilisation d'enzymes de restriction, et la synthèse chimique sur synthétiseur automatique. Les sondes et amorces susceptibles de s'hybrider dans des conditions de stringence déterminées à une séquence nucléotidique d'ADN ou d'ARN ou à un fragment nucléotidique tel que défini précédemment font partie de cette définition. Il est à la portée de l'homme du métier de définir les conditions de stringence appropriées. Des conditions de stringence caractéristiques sont celles qui correspondent à une combinaison de la température et de la concentration saline choisie approximativement entre 12 à 20°C sous le Tm (" melting temperature ") de l'hybride à l'étude. On peut ainsi se référer à l'ouvrage de George H. Keller et Mark M. Manak, DNA PROBES, second edition, Stockton Press, 1993, 49 West 24th St., New York, N.Y. 10010 USA. Les conditions de stringence pour discriminer même une seule mutation ponctuelle dans une séquence nucléique sont connues depuis au moins les années 1979 ; On peut citer à titre d'exemples Wallace R. B et al., DNA. Nucleic. Acids. Res. 6, 3543-3557 (1979), Wallace R. B et al., Science, 209, 1396-1400 (1980), Itakura K. and Riggs A.D., Science, 209, 1401-1405 (1980), Suggs S.V. et al., PNAS, 78, 6613-6617 (1981), Wallace R.B et al. DNA. Nucleic. Acids. Res., 9, 3647-3656 (1981), Wallace R.B et al. DNA. Nucleic. Acids. Res., 9, 879-894 (1981) et Conner B.J. etal, PNAS, 80, 278-282 (1983). Par ailleurs, on connaît des techniques pour la production d'anticorps anti-acides nucléiques. On peut citer à titre d'exemples Philippe Cros et al., Nucleic Acides Researc, 1994, Vol. 22, N°. 15, 2951-2957 ; Anderson, W.F. et al. (1988) Bioessays, 8 (2), 69-74 ; Lee, J.S. et al. (1984) FEBS Lett., 168, 303-306 ; Malfoy, B. et al. (1982)

Biochemistry, 21(22), 5463-5467 ; Stollar, B.D. et al., J.J. (eds) Methods in Enzymology, Academic Press, pp 70-85 ; Traincard, F. et al. (1989) J. Immunol. Meth., 123, 83-91 et Traincard, F. et al. (1989) Mol. Cell. Probes, 3, 27-38).

L'invention se rapporte aussi à :

- une composition vaccinale comprenant une séquence d'ADN codant pour au moins un polypeptide ou un fragment peptidique de l'invention, ledit ADN étant mélangé avec un véhicule et/ou un diluant et/ou un excipient approprié et pharmaceutiquement acceptable ;

- un oligonucléotide anti-sens ou anti-gène, caractérisé en ce qu'il est capable d'interférer spécifiquement avec la synthèse d'au moins un polypeptide ou un fragment peptidique de l'invention ;

- une composition pharmaceutique, caractérisée en ce qu'elle comprend au moins un oligonucléotide anti-sens ou un oligonucléotide anti-gène ;

- un vecteur, caractérisé en ce qu'il comprend au moins un gène d'intérêt thérapeutique ou vaccinal, ledit gène codant notamment

- (i) soit au moins pour un polypeptide ou un fragment peptidique de l'invention;

- (ii) soit au moins pour tout ou partie d'un anticorps capable de se fixer à au moins un polypeptide ou fragment peptidique défini en (i) ;

- (iii) soit au moins pour une molécule inhibitrice d'au moins un polypeptide ou fragment peptidique défini en (i) ;

- (iv) soit au moins pour un ligand ou toute partie d'un ligand capable de se fixer à au moins un polypeptide ou fragment peptidique défini en (i) et/ou d'inhiber sa fonction ;

- une composition thérapeutique ou vaccinale, caractérisée en ce qu'elle comprend, entre autre, un vecteur tel que défini ci-dessus et en ce que ledit gène

d'intérêt est placé sous la dépendance d'éléments assurant son expression *in vivo* ;

- un matériel biologique pour la préparation d'une composition pharmaceutique ou vaccinale, comprenant au moins une cellule, notamment une cellule ne produisant pas naturellement des anticorps, sous une forme permettant son administration dans un organisme mammifère, humain ou animal, ainsi qu'éventuellement sa culture préalable, ladite cellule étant génétiquement modifiée *in vitro* par au moins une séquence d'acide nucléique ou par au moins un fragment nucléotidique ou par au moins une molécule d'ADN ou par au moins un vecteur de l'invention, lesdits séquence d'acide nucléique, fragment nucléotidique, molécule d'ADN et gène du vecteur codant *in vivo* pour au moins un polypeptide ou un fragment peptidique de l'invention ou codant pour au moins tout ou partie d'un anticorps qui est capable de se lier à un polypeptide ou fragment peptidique de l'invention ou codant pour au moins une molécule inhibitrice de la fonction et/ou de la fixation et/ou de l'expression d'au moins un polypeptide ou d'un fragment peptidique ;

- une composition thérapeutique ou vaccinale comprenant ledit matériel biologique ;

- une cellule génétiquement modifiée, en particulier choisie parmi les cellules d'eucaryotes, telles que les cellules COS, CHO, Vero, BHK, PK 15, RK 13 ; les lignées cellulaires humaines de l'ostéosarcorme, les lignées cellulaires humaines HeLa et les lignées cellulaires humaines de l'hépatome, les lignées cellulaires d'insecte ; les cellules d'eucaryotes inférieurs, telles que les cellules de levure, en particulier les cellules issues de *Saccharomyces, Schizosaccharomyces, Kluveromyces, Hanseluna, Yarowia, Schwaniomyces, Zygosaccharomyces* et *Pichia*, et de préférence choisi parmi les cellules de *Saccharomyces cerevisiae, Saccharomyces carlsbergensis, Schizosaccharomyces pombe, Kluveromyces*

lactis et de *Pichia pastoris* ; les cellules de procaryotes, telles que celles issues de *E. coli* ; lesdites cellules étant transformées par au moins une séquence d'acide nucléique ou par au moins un fragment nucléotidique ou par une molécule d'ADN ou par un vecteur de l'invention; et

- une composition pharmaceutique ou vaccinale comprenant une telle cellule.

Les compositions pharmaceutiques définies ci-dessus sont des compositions vaccinales à ADN particulièrement avantageuses, en particulier par rapport aux compositions vaccinales " classiques " à base de protéine recombinante. En effet, l'utilisation à visée vaccinale de protéines recombinantes est un système lourd et onéreux, notamment parce qu'il exige de très importantes étapes de purification des antigènes recombinants. De plus, une des difficultés rencontrées est d'obtenir une persistance du vaccin suffisamment longue pour maintenir une bonne mémoire immunitaire. Au contraire, la méthode de vaccination par l'ADN, dont les avantages sont inhérents aux propriétés intrinsèques de l'ADN, est simple et peu coûteuse et est effectuée simplement par injection intramusculaire ou intradermique. De plus, il convient de noter que :

- les vaccins à ADN sont non infectieux/non réplicatifs,

- que du fait que l'immunisation par l'ADN est une forme de transfection *in vivo*, l'antigène viral est exprimé dans les cellules mammifères sous sa conformation native,

- comme dans le cas d'une infection virale, une large réponse immune, à la fois humorale et cellulaire est induite, et que

- de plus, les vaccins à ADN peuvent facilement être combinés en raison de leur homogénéité physico-chimique.

Enfin, l'invention a pour objet un procédé d'évaluation d'un agent thérapeutique selon lequel on administre à un animal des doses déterminées, en une dose ou en des doses répétées et à des intervalles de temps déterminés, au moins un polypeptide ou un fragment peptidique de l'invention, naturel, recombinant ou de synthèse, ou encore obtenu à partir d'un échantillon biologique éventuellement après un traitement préalable dudit échantillon biologique infecté par le virus HXHV, on prélève un échantillon biologique de l'animal, de préférence du sang ou du sérum et on réalise :

- (i) un dosage d'anticorps spécifique(s) du polypeptide ou du fragment polypeptidique ; et/ou
- (ii) un dosage de la réponse immune cellulaire induite contre le polypeptide ou le fragment polypeptidique, par exemple par un test d'activation *in vitro* de cellules lymphocytes T " helper " spécifique(s) du polypeptide ou du fragment polypeptidique.

Figure :

La figure représente le séquençage partiel de la bande d'environ 1,3 Kb. Dans la figure, le positionnement du fragment d'environ 200 paires de bases non séquencé est représenté par les symboles (-).Dans la figure, les fragments nucléotidiques indiqués en gras correspondent à des fragments nucléotidiques présentant une homologie ou identité de séquence avec SEQ ID NO : 1 de 100%. Leur positionnement respectifs par rapport à SEQ ID NO : 1 sont les suivants : 253-233, 254-273, 273-254.

Exemples

Exemple 1 : Extraction et extension

Les acides nucléiques ont été extraits à partir de 140 µl d'un échantillon de sérum d'un patient caractérisé comme étant HXHV positif par amplifications par PCR (Polymerase Chain Reaction), comme décrit dans la demande

de brevet PCT/FR02/04578, en utilisant le kit QIAamp Viral mini spin Kit (nom commercial) de la société Qiagen, en suivant le protocole préconisé par le fournisseur.

Une amorce biotinylée (Comp S6M13-biotin), dont la séquence est représentée ci-dessous, a ensuite été utilisée pour allonger la séquence SEQ ID NO : 1 d'intérêt. L'amorce biotinylée anti-sens utilisée correspond aux nucléotides 494-475 de SEQ ID NO : 1.

amorce anti-sens Comp S6M13:

5'-GCACTGCCGAGTTACATGGC-3' (SEQ ID NO :)

Pour l'extension, le kit GENEAmp XL PCR Kit (nom commercial) de la société Roche a été utilisé en respectant le protocole préconisé par le fournisseur.

La composition du mélange réactionnel de 50 µl est la suivante :

25 mM $Mg(OAC)_2$	2,4 µl
dNTPs 2,5 mM de chaque	4,0 µl
amorce Comp S6M13	2,0 µl (20 pico moles)
3.3X XL Buffer II	15,1 µl
rTth ADN polymerase (2 u/µl)	0,5 µl (1 u)
Matrice ADN	10 µl
Eau distillée	16 µl

L'extension a été réalisée selon le programme suivant :

Le mélange réactionnel a été chauffé à 92°C pendant 2 minutes et soumis ensuite à 35 cycles, chaque cycle comprenant un chauffage à 92°C pendant 30 secondes, un chauffage à 55°C pendant 30 secondes et un chauffage à 68°C pendant 3 minutes. L'extension finale a été réalisée par chauffage à 68°C pendant 10 minutes suivi d'un refroidissement à 4°C.

Exemple 2 : Capture de l'ADN double brin étendu.

Le produit d'extension obtenu selon le protocole décrit dans l'exemple 1 a été isolé en utilisant le kit Dynabeads Kilobase BINDER (nom commercial) de la société

Dynal, en suivant les instructions du fournisseur. Les billes (5 µl) ont premièrement été lavées deux fois dans le tampon Binding Buffer et resuspendues dans 20 µl de ce tampon. Un aliquot de 20 µl du produit d'extension a été ajouté et incubé pendant 3 heures à température ambiante sur un rouleau pour conserver les billes en suspension. L'ADN double brin a été purifié par deux lavages avec un tampon de lavage et un lavage à l'eau distillée et les billes ont ensuite été resuspendues dans 20 µl d'esu distillée et conservée à 4°C.

Exemple 3 : Digestion et circularisation.

5 µl de l'ADN double brin, capturé selon l'exemple 2, ont été digérés par l'enzyme *BsaWI* (NEB), dont le site de clivage correspondait à la position 299 de SEQ ID NO : 1, par chauffage à 60°C pendant 2 heures. L'enzyme a ensuite été inactivée par chauffage à 80°C pendant 20 minutes. Après quoi, le tube a été refroidi lentement et l'ADN digéré a été purifié en utilisant le kit QIA quick PCR purification Kit (nom commercial) de la société Qiagen. L'ADN purifié a ensuite été soumis a ligation à 4°C pendant une nuit en utilisant la ligase T4 commercialisée par la société Roche et la ligation a été achevée par chauffage à 65°C pendant 10 minutes.

Exemple 4 : Amplification.

10 µl du produit de ligation obtenu selon l'exemple 3 ont été utilisés comme matrice pour réaliser une PCR semi-nichée en utilisant le kit GeneAmp XL PCR Kit (nom commercial) de la société Roche. Les deux tours de PCR ont été réalisés de la même la façon, selon le protocole suivant :

Le mélange réactionnel a été chauffé à 94°C pendant 2 minutes et soumis ensuite à 35 cycles, chaque cycle comprenant un chauffage à 94°C pendant 30 secondes, un chauffage à 47°C pendant 30 secondes et un chauffage à

68°C pendant 3 minutes. Le mélange réactionnel a ensuite été soumis à un chauffage à 68°C pendant 10 minutes suivi d'un refroidissement à 4°C.

Premier tour de PCR :

Composition du mélange réactionnel (50 µl) :

25 mM $Mg(OAC)_2$	2,4 µl
dNTPs 2,5 mM de chaque	4,0 µl
amorce 1M13 sens (25µM)	1,0 µl
amorce CIRC 1 anti-sens (25 µM)	1,0 µl
3.3X XL Buffer II	15,1 µl
rTth ADN polymerase (2 u/µl)	0,5 µl (1 u)
Matrice ADN	10 µl
Eau	16 µl

Les couples d'amorces suivants ont été utilisés :

Amorce sens (1M13):

5'-CCCGCCCCGCTGATGAAAAG-3' (SEQ ID NO : 31)

Amorce anti-sens (CIRC 1)

5'-GCGATGGTTGAGTCTCGACTA-3' (SEQ ID NO: 32)

Deuxième tour de PCR :

Composition du mélange réactionnel (50 µl) :

25 mM $Mg(OAC)_2$	2,4 µl
dNTPs 2,5 mM de chaque	4,0 µl
amorce 1M13 sens (25µM)	1,0 µl
amorce 6BRACE5' anti-sens (25 µM)	1,0 µl
3.3X XL Buffer II	15,1 µl
rTth ADN polymerase (2 u/µl)	0,5 µl (1 u)
Produit du 1er tour	10 µl
Eau	16 µl

Les couples d'amorces suivants ont été utilisés :

Amorce sens (1M13):

5'-CCCGCCCCGCTGATGAAAAG-3' (SEQ ID NO : 31)

Amorce antisens (6BRACE5'):

5'-AGGTAGCAGGCGATATC-3' (SEQ ID NO: 33)

Les localisations des amorces dans la séquence XH (SEQ ID NO : 1) sont respectivement les suivantes :

1M13 : 254-273

CIR 1 : 253-233

6BRACE5' 94-77

Exemple 5 : Electrophorèse sur gel d'agarose et hybridation.

Les produits d'amplification obtenus selon l'exemple 4 ont été séparés par électrophorèse sur gel d'agarose 1,5%. Trois bandes dont les tailles étaient comprises entre 1,2 Kb et 2,5 Kb ont été observées sur le gel.

Les produits amplifiés ont été transférés sur une membrane de nylon Hybond-N+ (nom commercial) (Amersham Biosciences UK limited). La membrane a été hybridée à 42°C pendant une nuit avec le fragment XH complet marqué à son extrémité 3' au ^{32}P (généré en utilisant le kit Ready to Go DNA Labelling beads (nom commercial) de la société Amersham Pharmacia Biotech Inc.). Les lavages suivants ont été réalisés à 65°C : 2X SSC, 15 minutes, deux fois ; 1X SSC, 15 minutes, deux fois ; 0,5X SSC, 15 minutes, deux fois. La membrane a été soumise à autoradiographie à -80°C pendant une nuit. Les trois bandes présentaient des signaux faibles sur le film-X après développement.

Exemple 6 : Clonage et séquençage.

Les trois bandes ont respectivement été clonées dans le vecteur pCR2.1-TOPO (Invitrogen). Les clones ont ensuite été criblés par hybridation sur colonies et identifiés en utilisant l'enzymze *EcoRI* (Gibco BRL). Les clones positifs ont été sélectionnés pour être séquencés.

Les résultats du séquençage ont mis en évidence un fragment de 1133 paires de bases. La recherche effectuée dans les banques des bases de données n'a montré aucune homologie de séquences significative. Le fragment de 1133 paires de bases est référencé dans l'identificateur de

séquences en SEQ ID NOs : 2 et 3. Il est également représenté à la figure.

Exemple 7 : Répétition

En utilisant le même produit de digestion et de circularisation décrit dans l'exemple 3, une nouvelle amplification a été réalisée selon le protocole décrit dans l'exemple 4, suivi d'une électrophorèse sur gel d'agarose et de la procédure d'hybridation décrites dans l'exemple 5. Dans cet essai, une seule bande dont la taille était d'environ 1,3 Kb a été observée sur le gel. Après clonage et séquençage, comme décrit dans l'exemple 6, un fragment de 1133 paires de bases correspondant au fragment décrit dans l'exemple 6 (SEQ ID NOs : 2, 3 et figure) a été obtenu.

La pertinence de ce fragment de 1133 paires de bases par rapport au virus HXHV a été vérifiée comme décrit ci-dessous.

Dû aux limitations inhérentes au séquençage utilisé, la séquence de la bande d'environ 1,3 Kb visualisée sur gel s'est révélée être incomplète. En effet, un fragment d'environ 1300 paires de bases était attendu. Aussi, les présents inventeurs ont alors réalisé, avec une nouvelle procédure, un séquençage complet de la bande d'environ 1,3 Kb, comme décrit ci-dessous. La partie non séquencée dans le séquençage initial qui correspond à un fragment d'environ 200 paires de bases est représenté, pour sa localisation, dans la figure par les symboles (-). Le premier fragment séquencé est représenté en SEQ ID NO : 2 et le deuxième fragment séquencé est représenté en SEQ ID NO : 3 dans l'identificateur de séquences.

Exemple 8 : Pertinence du fragment de 1133 paires de bases.

Pour vérifier la pertinence du fragment de 1133 paires de bases par rapport au virus HXHV, des PCR nichées ont été réalisées en parallèle.

• A partir de fractions obtenues sur gradient de sucrose de 17 sérums, dont 10 étaient positifs pour l'ORF4 du virus HXHV décrite dans la demande de brevet PCT/FR02/04578 et 7 étaient négatifs pour cette même ORF4, les acides nucléiques ont été extraits et des PCR nichées ont été effectuées selon le protocole suivant en utilisant la Taq ADN polymérase de la société Promega :

Le mélange réactionnel a été chauffé à 94°C pendant 5 minutes et soumis ensuite à 35 cycles, chaque cycle comprenant un chauffage à 94°C pendant 45 secondes, un chauffage à 43°C pendant 45 secondes et un chauffage à 72°C pendant 1 minute. Le mélange réactionnel a ensuite été soumis à un chauffage à 72°C pendant 10 minutes suivi d'un refroidissement à 4°C.

Premier tour de PCR :

Composition du mélange réactionnel (50 µl) :

Tampon Taq avec MgCl2 10X	5,0 µl
dNTPs 10 mM de chaque	2,0 µl
amorce XF4 sens (25µM)	1,0 µl
amorce XB12 anti-sens (25 µM)	1,0 µl
Taq ADN polymerase (5 u/µl)	0,5 µl
Matrice ADN	10 µl
Eau	30,5 µl

Les couples d'amorces suivants ont été utilisés :

Amorce sens (XF4):

5' CCTTCTGGAGAGGGATTTC 3' (SEQ ID NO : 34)

Amorce anti-sens (XB12)

5' TGTTACCTGCTACTTCGTGC 3' (SEQ ID NO: 35)

Deuxième tour de PCR :

Composition du mélange réactionnel (50 µl) :

Tampon Taq avec MgCl2 10X	5,0 µl
dNTPs 10 mM de chaque	2,0 µl
amorce XF1 sens (25µM)	1,0 µl
amorce XB1 anti-sens (25 µM)	1,0 µl
Taq ADN polymerase (5 u/µl)	0,5 µl
Produit du 1er tour	10 µl
Eau	35,5 µl

Les couples d'amorces suivants ont été utilisés :

Amorce sens (XF1):

5' TAGAGTTGCGAGGCGTGACC 3' (SEQ ID NO : 36)

Amorce antisens (XB1):

5' CCTTATCCAGTGGCTTTTGGC 3' (SEQ ID NO: 37)

Les localisations des amorces dans la séquence SEQ ID NO : 4 sont respectivement les suivantes :

XF4: 482-500

XB12: 1255-1236

XF1: 944-963

XB1: 1186-1166

Les produits d'amplification obtenus selon l'exemple 4 ont été séparés par électrophorèse sur gel d'agarose 1,5%. Les produits amplifiés ont été transférés sur une membrane de nylon Hybond-N+ (nom commercial) (Amersham Biosciences UK limited). La membrane a été hybridée à 42°C pendant une nuit avec le produit du 2ème tour de PCR marqué à son extrémité 3' au ^{32}P. Le produit du tour 2 a été purifié avec le kit Qiaqick Gel Extraction Kit (nom commercial) et marqué en utilisant le kit Ready to Go DNA Labelling beads (nom commercial) de la société Amersham Pharmacia Biotech Inc.) Les lavages suivants ont été réalisés à 65°C : 2X SSC, 15 minutes, deux fois ; 1X SSC, 15 minutes, deux fois ; 0,5X SSC, 15 minutes, deux fois.

La membrane a été soumise à autoradiographie à -80°C pendant une nuit.

La bande de taille attendue d'environ 240 paires de bases est retrouvée dans les acides nucléiques amplifiés de 3 fractions sur les 10 qui étaient positives pour l'ORF4 du virus HXHV. Aucune bande n'a été révélée pour les 7 fractions qui étaient négatives pour l'ORF4 du virus HXHV.

• Les acides nucléiques extraits de 15 sérums de patients Non A-E, dont 9 étaient positifs pour l'ORF4 de HXHV et 6 étaient négatifs pour cette même ORF ont été amplifiés par PCR nichée avec le même protocole que celui décrit ci-dessus. Les produits amplifiés obtenus ont ensuite été séparés par électrophorèse sur gel d'agarose, les membranes ont été hybridées et les bandes ont été révélées par autoradiographie selon le même protocole que celui décrit ci-dessus.

La bande de taille attendue d'environ 240 paires de bases est retrouvée dans les acides nucléiques amplifiés de 3 sérums sur les 9 qui étaient positifs pour l'ORF4 du virus HXHV. Aucune bande n'a été révélée pour les 6 sérums qui étaient négatifs pour l'ORF4 du virus HXHV.

Les résultats obtenus à partir de fractions de sérums et de sérums confirment donc que la séquence de 1133 paires de bases est associée au virus HXHV.

Exemple 9 : Séquençage complet de la bande d'environ 1,3 Kb.

Les produits PCR ont été purifiés par digestion enzymatique (Enzyme Exosup - nom commercial). La quantification des acides nucléiques a été réalisée par dosage fluorométrique. La réaction de séquençage a été réalisée grâce à une réaction enzymatique en présence d'une amorce spécifique de la région à séquencer. Les produits ont ensuite été injectés dans le séquenceur Apply Biosystem 3730 XL (nom commercial). La séquence ADN

obtenue est une séquence de 1314 paires de bases représentée en SEQ ID NO : 4.

REVENDICATIONS

1. Séquence d'acide nucléique isolée susceptible d'être obtenue à partir du génome du virus HXHV, ladite séquence d'acide nucléique comprenant la séquence SEQ ID NO : 4 ou la séquence complémentaire de SEQ ID NO : 4.

2. Séquence d'acide nucléique isolée susceptible d'être obtenue à partir du génome du virus HXHV, ladite séquence d'acide nucléique consistant en la séquence SEQ ID NO : 4 ou en la séquence complémentaire de SEQ ID NO : 4.

3. Fragment nucléotidique d'ADN, caractérisé en ce qu'il comprend ou consiste en une séquence nucléotidique d'au moins 12 nucléotides contigus appartenant à SEQ ID NO : 4 ou à son complémentaire.

4. Fragment selon la revendication 3, caractérisé en ce qu'il comprend ou consiste en une séquence d'au moins 15 nucléotides contigus appartenant à SEQ ID NO : 4 ou à son complémentaire.

5. Fragment selon la revendication 3 ou 4, caractérisé en ce qu'il comprend ou consiste en une séquence d'au moins 18 nucléotides contigus appartenant à SEQ ID NO : 4 ou à son complémentaire.

6. Fragment selon l'une quelconque des revendications 3 à 5, caractérisé en ce qu'il comprend ou consiste en une séquence d'au moins 20, 21, 22, 23, 24, 27, 30, 33,36, 39, 42, 45, 48 51 ou 54 nucléotides contigus appartenant à SEQ ID NO : 4 ou à son complémentaire.

7. Fragment nucléotidique d'ADN, caractérisé en ce qu'il comprend ou consiste en une séquence nucléotidique qui sur au moins 12 nucléotides contigus présente au moins 90% d'identité, de préférence au moins 95% d'identité et avantageusement au moins 98% ou 99% d'identité avec la SEQ ID NO : 4 ou avec la séquence complémentaire de SEQ ID NO : 4, à l'exclusion des séquences TAGTCGAGACTCAACCATCGC

et CCCGCCCCGCTGATGAAAAG et des séquences nucléotidiques complémentaires desdites séquences.

8. Fragment nucléotidique d'ADN, caractérisé en ce qu'il comprend ou consiste en une séquence nucléotidique qui sur au moins 15 nucléotides contigus présente au moins 90% d'identité, de préférence au moins 95% d'identité et avantageusement au moins 98% ou 99% d'identité avec la SEQ ID NO : 4 ou avec la séquence complémentaire de SEQ ID NO : 4, à l'exclusion des séquences TAGTCGAGACTCAACCATCGC et CCCGCCCCGCTGATGAAAAG, et des séquences nucléotidiques complémentaires desdites séquences.

9. Fragment nucléotidique d'ADN, caractérisé en ce qu'il comprend ou consiste en une séquence nucléotidique qui sur au moins 18 nucléotides contigus présente au moins 90% d'identité, de préférence au moins 95% d'identité et avantageusement au moins 98% ou 99% d'identité avec la SEQ ID NO : 4 ou avec la séquence complémentaire de SEQ ID NO : 4, à l'exclusion des séquences TAGTCGAGACTCAACCATCGC et CCCGCCCCGCTGATGAAAAG, et des séquences nucléotidiques complémentaires desdites séquences.

10. Fragment nucléotidique d'ADN, caractérisé en ce qu'il comprend ou consiste en une séquence nucléotidique qui sur au moins 20, 21, 22, 23, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51 ou 54 nucléotides contigus présente au moins 90% d'identité, de préférence au moins 95% d'identité et avantageusement au moins 98% ou 99%d'identité avec la SEQ ID NO : 4 ou avec la séquence complémentaire de SEQ ID NO : 4, à l'exclusion des séquences TAGTCGAGACTCAACCATCGC et CCCGCCCCGCTGATGAAAAG, et des séquences nucléotidiques complémentaires desdites séquences.

11. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 2 et se terminant au nucléotide 286 de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

12. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 4 et se terminant au nucléotide 144 de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

13. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 180 et se terminant au nucléotide 1004 de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

14. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 614 et se terminant au nucléotide 820 de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

15. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 1228 et se terminant au nucléotide 1314 de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

16. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 1283 et se terminant au nucléotide 1197 de la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

17. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 1264 et se terminant au nucléotide 1067 de la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

18. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 1209 et se terminant au nucléotide 1099 de

la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

19. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 819 et se terminant au nucléotide 736 de la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

20. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 800 et se terminant au nucléotide 6 de la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

21. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 784 et se terminant au nucléotide 629 de la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

22. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 610 et se terminant au nucléotide 410 de la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

23. Fragment selon l'une quelconque des revendications 3 à 10, caractérisé en ce que lesdits nucléotides contigus appartiennent au segment commençant au nucléotide 391 et se terminant au nucléotide 221 de la séquence complémentaire de SEQ ID NO : 4 ou fragment complémentaire dudit fragment.

24. Fragment selon l'une quelconque des revendications 3 à 6, caractérisé en ce qu'il comprend ou consiste en l'une quelconque des séquences SEQ ID NOs 5 à

17 ou l'une quelconque des séquences complémentaires des séquences SEQ ID Nos 5 à 17.

25. Produit de transcription d'une séquence telle que définie dans l'une quelconque des revendications 1 et 2 ou d'un fragment tel que défini dans l'une quelconque des revendications 3 à 24.

26. Molécule d'ADN, caractérisée en ce qu'elle comprend ou consiste en une séquence telle que définie dans l'une quelconque des revendications 1 et 2 ou un fragment tel que défini à l'une quelconque des revendications 3 à 24.

27. Molécule d'ARN, caractérisée en ce qu'elle comprend ou consiste en un produit de transcription d'une molécule d'ADN telle que définie à la revendication 26.

28. Polypeptide dont la séquence polypeptidique est codée par une séquence telle que définie dans l'une quelconque des revendications 1 et 2 ou par un fragment tel que défini dans l'une quelconque des revendications 3 à 24.

29. Polypeptide selon la revendication 28, dont la séquence polypeptidique comprend ou consiste en l'une quelconque des séquences SEQ ID NO : 18 à 30 ou en une séquence polypeptidique équivalente à l'une quelconque des séquences SEQ ID NO : 18 à 30 dans laquelle (i) les acides aminés alanine, proline, glycine sont des équivalents, (ii) les acides aminés acide aspartique, acide glutamique sont des équivalents, (iii) les acides aminés histidine, lysine, arginine sont des équivalents, (iv) les acides aminés asparagine, glutamine, sérine, thréonine sont des équivalents, (v) les acides aminés phénylalanine, tyrosine, tryptophane sont des équivalents et (vi) les acides aminés isoleucine, leucine, valine, méthionine sont des équivalents.

30. Fragment polypeptidique selon la revendication 28, comprenant ou consistant en une séquence peptidique d'au moins 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

17 ou acides aminés appartenant à l'une quelconque des séquences SEQ ID NOs : 18 à 30 ou à une séquence équivalente à l'une quelconque des séquences SEQ ID NO : 18 à 30 dans laquelle (i) les acides aminés alanine, proline, glycine sont des équivalents, (ii) les acides aminés acide aspartique, acide glutamique sont des équivalents, (iii) les acides aminés histidine, lysine, arginine sont des équivalents, (iv) les acides aminés asparagine, glutamine, sérine, thréonine sont des équivalents, (v) les acides aminés phénylalanine, tyrosine, tryptophane sont des équivalents et (vi) les acides aminés isoleucine, leucine, valine, méthionine sont des équivalents.

31. Epitope caractérisé en ce qu'il comprend ou consiste en une séquence peptidique d'au moins 6, 8, 9, 10, 12, 15 ou 18 acides aminés et au plus de 10, 12, 15 ou 18 acides aminés, en particulier en ce que sa séquence consiste en une séquence peptidique de 6 à 10 acides aminés, de 6 à 12 acides aminés, de 6 à 15 acides aminés, de 6 à 18 acides aminés, de 8 à 10 acides aminés, de 8 à 12 acides aminés, de 8 à 15 acides aminés, de 8 à 18 acides aminés et de 15 à 18 acides aminés de l'une quelconque des séquences représentées en SEQ ID NO : 18 à 30 ou d'une séquence polypeptidique fonctionnellement équivalente auxdites séquences SEQ ID NO : 18 à 30.

32. Cassette d'expression fonctionnelle dans une cellule issue d'un organisme procaryote ou eucaryote, permettant l'expression d'une séquence d'acide nucléique selon l'une quelconque des revendications 1 ou 2, ou d'un fragment selon l'une quelconque des revendications 3 à 24, ou d'une molécule d'ADN selon la revendications 26, placée sous le contrôle des éléments nécessaires à son expression.

33. Vecteur comprenant une cassette d'expression selon la revendication 32.

34. Cellule issue d'un organisme eucaryote ou procaryote comprenant une cassette d'expression selon la revendication 32 ou un vecteur d'expression selon la revendication 33.

35. Cellule selon la revendication 34, caractérisée en ce qu'elle est issue d'un organisme eucaryote, en particulier les cellules provenant d'animaux tels que les mammifères, les reptiles, les insectes, de préférence les cellules choisies parmi les cellules COS, CHO, Vero, BHK, PK 15, RK 13 ; les lignées cellulaires humaines de l'ostéosarcorme, les lignées cellulaires humaines HeLa et les lignées cellulaires humaines de l'hépatome ; les lignées cellulaires d'insecte.

36. Cellule selon la revendication 34, caractérisée en ce qu'elle est issue d'un organisme eucaryote inférieur, en particulier issue de levures, telles que *Saccharomyces, Schizosaccharomyces, Kluveromyces, Hanseluna, Yarowia, Schwaniomyces, Zygosaccharomyces* et *Pichia*, et de préférence choisi parmi les cellules de *Saccharomyces cerevisiae, Saccharomyces carlsbergensis, Schizosaccharomyces pombe, Kluveromyces lactis* et de *Pichia pastoris*.

37. Cellule selon la revendication 34, caractérisée en ce qu'elle est issue d'un organisme procaryote de préférence *E. coli*.

38. Polypeptide susceptible d'être produit par une cassette d'expression selon la revendication 32, un vecteur selon la revendication 33, ou une cellule selon l'une quelconque des revendications 34 à 37.

39. Procédé pour préparer un polypeptide selon la revendication 28 ou 29 ou un fragment peptidique selon la revendication 30, selon lequel on cultive une cellule hôte telle que définie à l'une quelconque des revendications 34 à 37 dans un milieu de culture approprié, et on purifie

ledit polypeptide ou ledit fragment peptidique produit jusqu'à un degré de pureté requis.

40. Polypeptide immunogène comprenant ou consistant en un polypeptide tel que défini à la revendication 28 ou 29 ou un fragment peptidique tel que défini à la revendication 30.

41. Anticorps monoclonal ou polyclonal susceptible d'être obtenu par immunisation d'un animal mammifère avec un polypeptide immunogène tel que défini dans la revendication 40.

42. Composition diagnostique, caractérisée en ce qu'elle comprend un polypeptide tel que défini à la revendication 28 ou 29 ou un fragment polypeptidique tel que défini à la revendication 30.

43. Composition diagnostique, caractérisée en ce qu'elle comprend un anticorps monoclonal ou un anticorps polyclonal tel que défini dans la revendication 41.

44. Procédé pour détecter des anticorps dirigés contre le virus HXHV ou au moins un polypeptide tel que défini à la revendication 28 ou 29 ou un fragment peptidique tel que défini dans la revendications 30, selon lequel on met en contact un échantillon biologique d'un patient suspecté être infecté par le virus HXHV avec une composition diagnostique telle que définie dans la revendication 42, dans des conditions prédéterminées qui permettent la formation de complexes anticorps/antigène et on détecte la formation desdits complexes.

45. Procédé pour détecter un polypeptide tel que défini à la revendication 28 ou 29 ou un fragment peptidique tel que défini à la revendication 30, dans un échantillon biologique d'un patient suspecté être infecté par le virus HXHV, selon lequel on met en contact l'échantillon biologique avec une composition diagnostique telle que définie dans la revendication 43, dans des conditions prédéterminées qui permettent la formation de

complexes anticorps/antigène et on détecte la formation desdits complexes.

46. Composition immunogène ou vaccinale, caractérisée en ce qu'elle comprend un polypeptide tel que défini à la revendication 28 ou 29 ou un fragment peptidique tel que défini dans la revendication 30, associé à un véhicule et/ou adjuvant et/ou diluant approprié et/ou à un excipient pharmaceutiquement acceptable.

47. Sonde d'au moins 12 nucléotides, caractérisée en ce qu'elle est susceptible de s'hybrider à une séquence d'acide nucléique telle que définie dans la revendication 1 ou 2, ou à un fragment nucléotidique tel que défini dans l'une quelconque des revendications 3 à 24, ou à une molécule d'ADN ou d'ARN telle que définie dans les revendications 26 ou 27, l'hybridation étant réalisée dans des conditions de stringence déterminées.

48. Amorce d'au moins 12 nucléotides, caractérisée en ce qu'elle est susceptible de s'hybrider à une séquence d'acide nucléique telle que définie dans la revendication 1 ou 2, ou à un fragment nucléotidique tel que défini dans l'une quelconque des revendications 3 à 24, ou à une molécule d'ADN ou d'ARN telle que définie dans la revendication 26 ou 27 l'hybridation étant réalisée dans des conditions de stringence déterminées.

49. Amorce selon la revendication 48, caractérisée en ce qu'elle est choisie parmi les amorces SEQ ID Nos 32 à 37.

50. Couple d'amorces selon la revendication 48, caractérisé en ce qu'il est choisi parmi un des couples suivants : SEQ ID NO : 31 / SEQ ID NO : 32, SEQ ID NO : 31 / SEQ ID NO : 33, SEQ ID NO : 34 / SEQ ID NO : 35, , SEQ ID NO : 36 / SEQ ID NO : 37.

51. Anticorps anti-acide nucléique, caractérisé en ce qu'il est susceptible de se lier à une séquence d'acide nucléique telle que définie dans la revendication 1 ou 2, ou à un fragment nucléotidique tel que défini dans l'une

quelconque des revendications 3 à 24, ou à une molécule d'ADN ou d'ARN telle que définie dans la revendication 26 ou 27.

52. Composition diagnostique, caractérisée en ce qu'elle comprend au moins une sonde ou au moins une amorce ou au moins un couple d'amorces ou un anticorps anti-acide nucléique tel(le) que défini(e) dans les revendications 47, 48, 49, 50 ou 51.

53. Procédé de détection d'un ADN ou d'un ARN viral, dans un échantillon biologique d'un patient suspecté être infecté par le virus HXHV, selon lequel on traite si nécessaire ledit échantillon pour en extraire l'ADN ou l'ARN, on met en contact ledit ADN ou ARN avec au moins une sonde ou avec au moins une amorce ou avec au moins un couple d'amorces tel(le) que défini(e) dans les revendications 47, 48, 49 ou 50, dans des conditions de stringence déterminées, et on détecte la présence d'ADN ou d'ARN viral dans l'échantillon soit par mise en évidence d'une hybridation dudit ADN ou ARN viral avec au moins une sonde telle que définie dans la revendication 47, soit par amplification dudit ADN ou ARN à l'aide d'au moins une amorce telle que définie dans la revendication 48 ou 49 ou d'au moins un couple d'amorces tel que défini dans la revendication 50.

54. Procédé de détection d'ADN et/ou d'ARN viral du virus HXHV, selon lequel on prélève un échantillon biologique, tel que du sérum, plasma ou sang d'un patient, on traite si nécessaire ledit échantillon pour en extraite l'ADN et/ou l'ARN, on met en contact ledit échantillon avec au moins un anticorps anti-acide nucléique tel que défini dans la revendication 51, ledit anticorps étant éventuellement marqué par tout marqueur approprié et on met en évidence la formation d'un complexe acide nucléique/anticorps.

55. Composition vaccinale comprenant une séquence d'ADN codant pour au moins un polypeptide tel que défini

dans la revendication 28 ou 29 ou codant pour au moins un fragment peptidique tel que défini dans la revendication 30, ledit ADN étant mélangé avec un véhicule et/ou un diluant et/ou un excipient pharmaceutiquement acceptable.

56. Vecteur comprenant au moins un gène d'intérêt thérapeutique ou vaccinal, ledit gène codant notamment au moins pour un polypeptide ou un fragment peptidique tel que défini dans l'une quelconque des revendications 28, 29 et 30.

57. Composition thérapeutique ou vaccinale, caractérisée en ce qu'elle comprend un vecteur tel que défini dans la revendication 56 et en ce que ledit gène d'intérêt est placé sous la dépendance d'éléments assurant son expression *in vivo*.

58. Cellule génétiquement modifiée, en particulier choisie parmi les cellules d'eucaryotes, telles que les cellules COS, CHO, Vero, BHK, PK 15, RK 13 ; les lignées cellulaires humaines de l'ostéosarcorme, les lignées cellulaires humaines HeLa et les lignées cellulaires humaines de l'hépatome, les lignées cellulaires d'insecte ; les cellules d'eucaryotes inférieurs, telles que les cellules de levure, en particulier les cellules issues de *Saccharomyces, Schizosaccharomyces, Kluveromyces, Hanseluna, Yarowia, Schwaniomyces, Zygosaccharomyces* et *Pichia*, et de préférence choisi parmi les cellules de *Saccharomyces cerevisiae, Saccharomyces carlsbergensis, Schizosaccharomyces pombe, Kluveromyces lactis* et de *Pichia pastoris* ; les cellules de procaryotes, telles que celles issues de *E. coli* ; lesdites cellules étant transformées par au moins une séquence d'acide nucléique selon la revendication 1 ou 2 ou par au moins un fragment nucléotidique selon l'une quelconque des revendications 3 à 24 ou par une molécule d'ADN selon la revendication 26 ou par un vecteur selon la revendication 56.

59. Composition pharmaceutique ou vaccinale, caractérisée en ce qu'elle comprend une cellule telle que définie dans la revendication 58.

5

Figure

TAGTCGAGACTCAACCATCGCT**CCCGCCCCGCTGATGAAAAG**GTCGCTCGGCTCAAGCGCGAACTGTC
GCGTGTTACCAAGGAACGAGATTTTTTACGAGACGCGGCAGCGTACTTCGCGAAGCAATCGCCGAACG
GTACGCGGTGATCGAGCGCTGCCGCAGCGACTACCCCATTGGGATGATGTGCCGCTGCCTTCAAGTGT
CGACCAGTGGGTTCTACGCCTGGGCCAGGCGAAAGCCGGGGCCGCGTGCCCAGGCGAATTCGCGTCTC
TTGGAGCGCATGCGTGAAATCCACGAGGACAGCCGAGGCATCATCGGCGCGCGTCGGATGCAGGAAGA
CCTCGCCGACGAAGGCATGCCCGCCAGCTTGAATCGGGTGGCCCGCGTCATGGCCAAGGCCGGGCTTC
AGGGCTGGCCGCGGCGAAAGAAGCGTGGCTTTCCGCGCAAGCCGCCGACGCGTCGTCCCGAGGGCGTC
AGGAACCTTCTGGAGAGGGATTTCTCGGCGCTCGAACCGGAGACGAAGTGGGTAACCGACATCACCGA
GATCGTCACCGACGAGGG----------
TACCAGAAGTTCCTCGGCAGCCATGCCTTGGTCTGCAGCATGAGCGAGGTCGGCCATTGCGGCGACAA
CGCAGCATGTGAGGGATTCTTCGGGCTACTGAAGCGAGAGTGGATCTACCAAACCCGCTACAGCACAA
GAAGGGAAGCTCGGGCCGACGTCTTTGCCTACCTGGAGCGGTTCCACGACCCGCGGATGCGCCGTAGA
GTTGCGAGGCGTGACCGGGAGTTTCAAGCCTTAATCAAACCGTCCGCGGAAACGGGGTAGAACCCGAG
TCCACTTACCGCCGGTGCGGCGCAGGTCGCCGCCCCACACCACGCAGGTTAAGTCGAGTTCCGAACCC
TGCACCTGAAACTCAGTGGCGACGTCTTCGAGGTAGTACGACGTCTTCGCGGTGATACAAGAAACGTG
TGTCCTCCTTGCCGGCCAAAAGCCACTGGATAAGGTCGACCTTTAAGCGCATTCCCATAGCGTACTGG
GACCCCTGATGCCGAGGCACGAAGTAGCAGGTAACATCGTGTCATGCACAAGCAATCGGATCATGTCG
TCTCGCT**CTTTTCATGAGCGGGGCGGG**

SEQUENCE LISTING

<110> BIOMERIEUX

INSERM (INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE)

<120> Séquences d'acides nucléiques et protéiques du virus HXHV et utilisations

<130> HXHV1

<160> 37

<170> PatentIn version 3.1

<210> 1

<211> 1362

<212> DNA

<213> virus

<400> 1

actaccaaca gatcctcgac gaactgcgcc aggaactggc cgagcactac ctgctgcgca 60

gcgacctggc gatccaggat atcgcctgct acctcggttt caccgagtca cgctcgttcc 120

accgcagttt caagagctgg accgggcaga cgccgggcga gtttcgcgag agccggcgcc 180

gggataatcc gctgggctag cgcgatatgg ccggaaacgc cgtgccagcc agtagtcgag 240

actcaaccat cgccccgccc cgctgatgaa aagcgccacg agcgcagcca cggccggcac 300

cggtgaggtt tgccaatggc atatcagtcc tcccggcgcc cttactcgtt cttatcgcca 360

ctgcacgtgc cttcaatacg ggagccttcc tgcgccttct cggcagcggt caggctgtag 420

ccgccggcca gttcctgctc agcgaagggg atgctagtgg cgtgggcagt gaacgccatg 480

taactcggca gtgcagcgcc ctagggtctg ttgccgtttc gcgcacggcc gcgtcgaaac 540

ggcaacagac cctaggtggc agtcagggta ttggcatctc tccatcggtt tcgaatacgg 600

cgccaggttg gcgccctcgc agcaatggac gaggcaggga tgcgggcgtt acagcgggcg 660

aaaaagattt ctcgtagccc gatgaaatac gggggcgctt tgctcgccag caatcgcggc 720

tacgactgca tggacgcagg aggtagagcg aagcaggatg vvagagcaga aagctctctc 780

ccacagacac agaaacatcc accgcacggt aggaggtgat tcaaatgatc aggcatctcc 840

tctggttgga ctgcatggcc gctgcgagca cgggcgttgt ggttctgttg ctggcccccc 900

tggttgagcg gctggtatgc cctgcccggc gagctgctga gcttcatcgg cgcgatcaat 960

atcgcctacg cctgcttttc catttcgctg gcgattcgcc tgcgacgcgc cgaagcgcta 1020

atcaagctgc tggcagtggc caacggactc tgggcgttgg catgccttgg catcgctacg 1080

atctttgccc cgctcatgac gctaccgggg ctttgtcatg tgctcggcga ggctgcatcc 1140

gtcgcaggcc tgggcatgct ggagtggaaa tggcgcaggc agctgctggt ggctggcgaa 1200

cagggcgttt cgactcagct tgtcgcggtc cagtaaccgt cacaggtatt caggcgaaga 1260

tccggcgatg gctgttcagg ccatcatcag ccctatcctt cagccctgtg aaagcggttc 1320

ttgcccgcgt gcttggccgc gtacctcggc cccgaccacg ct 1362

<210> 2

<211> 562

<212> DNA

<213> virus

<400> 2

tagtcgagac tcaaccatcg ctcccgcccc gctgatgaaa aggtcgctcg gctcaagcgc 60

gaactgtcgc gtgttaccaa ggaacgagat tttttacgag acgcggcagc gtacttcgcg 120

aagcaatcgc cgaacggtac gcggtgatcg agcgctgccg cagcgactac cccattggga 180

tgatgtgccg ctgccttcaa gtgtcgacca gtgggttcta cgcctgggcc aggcgaaagc 240

cggggccgcg tgcccaggcg aattcgcgtc tcttggagcg catgcgtgaa atccacgagg 300

acagccgagg catcatcggc gcgcgtcgga tgcaggaaga cctcgccgac gaaggcatgc 360

ccgccagctt gaatcgggtg gcccgcgtca tggccaaggc cgggcttcag ggctggccgc 420

ggcgaaagaa gcgtggcttt ccgcgcaagc cgccgacgcg tcgtcccgag ggcgtcagga 480

accttctgga gagggatttc tcggcgctcg aaccggagac gaagtgggta accgacatca 540

ccgagatcgt caccgacgag gg 562

<210> 3

<211> 571

<212> DNA

<213> virus

<400> 3
taccagaagt tcctcggcag ccatgccttg gtctgcagca tgagcgaggt cggccattgc 60

ggcgacaacg cagcatgtga gggattcttc gggctactga agcgagagtg gatctaccaa 120

acccgctaca gcacaagaag ggaagctcgg gccgacgtct ttgcctacct ggagcggttc 180

cacgacccgc ggatgcgccg tagagttgcg aggcgtgacc gggagtttca agccttaatc 240

aaaccgtccg cggaaacggg gtagaacccg agtccactta ccgccggtgc ggcgcaggtc 300

gccgccccac accacgcagg ttaagtcgag ttccgaaccc tgcacctgaa actcagtggc 360

gacgtcttcg aggtagtacg acgtcttcgc ggtgatacaa gaaacgtgtg tcctccttgc 420

cggccaaaag ccactggata aggtcgacct ttaagcgcat tcccatagcg tactgggacc 480

cctgatgccg aggcacgaag tagcaggtaa catcgtgtca tgcacaagca atcggatcat 540

gtcgtctcgc tcttttcatg agcggggcgg g 571

<210> 4

<211> 1314

<212> DNA

<213> virus

<400> 4
tagtcgagac tcaaccatcg ctcccgcccc gctgatgaaa aggtcgctcg gctcaagcgc 60

gaactgtcgc gtgttaccaa ggaacgagat tttttacgag acgcggcagc gtacttcgcg 120

aagcaatcgc cgaacggtac gcggtgatcg agcgctgccg cagcgactac cccattggga 180

tgatgtgccg ctgccttcaa gtgtcgacca gtgggttcta cgcctgggcc aggcgaaagc 240

cggggccgcg tgcccaggcg aattcgcgtc tcttggagcg catgcgtgaa atccacgagg 300

acagccgagg catcatcggc gcgcgtcgga tgcaggaaga cctcgccgac gaaggcatgc 360

ccgccagctt gaatcgggtg gcccgcgtca tggccaaggc cgggcttcag ggctggccgc 420

ggcgaaagaa gcgtggcttt ccgcgcaagc cgccgacgcg tcgtcccgag ggcgtcagga 480

accttctgga gagggatttc tcggcgctcg aaccggagac gaagtgggta accgacatca 540

ccgagatcgt caccgacgag ggaaaactcc atctctgcgt cgtcctcgac ctgtacagca 600

aactcatcat gggatggtcg atgcatcacc ggcaggatcg ccacatggtg gttcgcgcgg 660

tacagatggc ggtttggcag cgcgagggcg gcgacgaggt gatcctgcat tccgatcgcg 720

gcgggcagtt catcagcgat acgtaccaga agttcctcgg cagccatgcc ttggtctgca 780

gcatgagcga ggtcggccat tgcggcgaca acgcagcatg tgagggattc ttcgggctac 840

tgaagcgaga gtggatctac caaacccgct acagcacaag aagggaagct cgggccgacg 900

tctttgccta cctggagcgg ttccacgacc cgcggatgcg ccgtagagtt gcgaggcgtg 960

accgggagtt tcaagcctta atcaaaccgt ccgcggaaac ggggtagaac ccgagtccac 1020

ttaccgccgg tgcggcgcag gtcgccgccc cacaccacgc aggttaagtc gagttccgaa 1080

ccctgcacct gaaactcagt ggcgacgtct tcgaggtagt acgacgtctt cgcggtgata 1140

caagaaacgt gtgtcctcct tgccggccaa aagccactgg ataaggtcga cctttaagcg 1200

cattcccata gcgtactggg acccctgatg ccgaggcacg aagtagcagg taacatcgtg 1260

tcatgcacaa gcaatcggat catgtcgtct cgctcttttc atgagcgggg cggg 1314

<210> 5

<211> 285

<212> DNA

<213> virus

<400> 5
agtcgagact caaccatcgc tcccgccccg ctgatgaaaa ggtcgctcgg ctcaagcgcg 60

aactgtcgcg tgttaccaag gaacgagatt ttttacgaga cgcggcagcg tacttcgcga 120

agcaatcgcc gaacggtacg cggtgatcga gcgctgccgc agcgactacc ccattgggat 180

gatgtgccgc tgccttcaag tgtcgaccag tgggttctac gcctgggcca ggcgaaagcc 240

ggggccgcgt gcccaggcga attcgcgtct cttggagcgc atgcg 285

<210> 6

<211> 141

<212> DNA

<213> virus

<400> 6
tcgagactca accatcgctc ccgccccgct gatgaaaagg tcgctcggct caagcgcgaa 60

ctgtcgcgtg ttaccaagga acgagatttt ttacgagacg cggcagcgta cttcgcgaag 120

caatcgccga acggtacgcg g 141

<210> 7

<211> 825

<212> DNA

<213> virus

<400> 7
atgatgtgcc gctgccttca agtgtcgacc agtgggttct acgcctgggc caggcgaaag 60

ccggggccgc gtgcccaggc gaattcgcgt ctcttggagc gcatgcgtga aatccacgag 120

gacagccgag gcatcatcgg cgcgcgtcgg atgcaggaag acctcgccga cgaaggcatg 180

cccgccagct tgaatcgggt ggcccgcgtc atggccaagg ccgggcttca gggctggccg 240

cggcgaaaga agcgtggctt tccgcgcaag ccgccgacgc gtcgtcccga gggcgtcagg 300

aaccttctgg agagggattt ctcggcgctc gaaccggaga cgaagtgggt aaccgacatc 360

accgagatcg tcaccgacga gggaaaactc catctctgcg tcgtcctcga cctgtacagc 420

aaactcatca tgggatggtc gatgcatcac cggcaggatc gccacatggt ggttcgcgcg 480

gtacagatgg cggtttggca gcgcgagggc ggcgacgagg tgatcctgca ttccgatcgc 540

ggcgggcagt tcatcagcga tacgtaccag aagttcctcg gcagccatgc cttggtctgc 600

agcatgagcg aggtcggcca ttgcggcgac aacgcagcat gtgagggatt cttcgggcta 660

ctgaagcgag agtggatcta ccaaacccgc tacagcacaa gaagggaagc tcgggccgac 720

gtctttgcct acctggagcg gttccacgac ccgcggatgc gccgtagagt tgcgaggcgt 780

gaccgggagt ttcaagcctt aatcaaaccg tccgcggaaa cgggg 825

<210> 8

<211> 207

<212> DNA

<213> virus

<400> 8
atggtcgatg catcaccggc aggatcgcca catggtggtt cgcgcggtac agatggcggt 60

ttggcagcgc gagggcggcg acgaggtgat cctgcattcc gatcgcggcg ggcagttcat 120

cagcgatacg taccagaagt tcctcggcag ccatgccttg gtctgcagca tgagcgaggt 180

cggccattgc ggcgacaacg cagcatg 207

<210> 9

<211> 87

<212> DNA

<213> virus

<400> 9
atgccgaggc acgaagtagc aggtaacatc gtgtcatgca caagcaatcg gatcatgtcg 60

tctcgctctt ttcatgagcg gggcggg 87

<210> 10

<211> 87

<212> DNA

<213> virus

<400> 10
agcgcattcc catagcgtac tgggacccct gatgccgagg cacgaagtag caggtaacat 60

cgtgtcatgc acaagcaatc ggatcat 87

<210> 11

<211> 198

<212> DNA

<213> virus

<400> 11
agtcgagttc cgaaccctgc acctgaaact cagtggcgac gtcttcgagg tagtacgacg 60

tcttcgcggt gatacaagaa acgtgtgtcc tccttgccgg ccaaaagcca ctggataagg 120

tcgaccttta agcgcattcc catagcgtac tgggacccct gatgccgagg cacgaagtag 180

caggtaacat cgtgtcat 198

<210> 12

<211> 111

<212> DNA

<213> virus

<400> 12
gtggcgacgt cttcgaggta gtacgacgtc ttcgcggtga tacaagaaac gtgtgtcctc 60

cttgccggcc aaaagccact ggataaggtc gacctttaag cgcattccca t 111

<210> 13

<211> 84

<212> DNA

<213> virus

<400> 13
gcgatacgta ccagaagttc ctcggcagcc atgccttggt ctgcagcatg agcgaggtcg 60

gccattgcgg cgacaacgca gcat 84

<210> 14

<211> 795

<212> DNA

<213> virus

<400> 14
gagactcaac catcgctccc gccccgctga tgaaaaggtc gctcggctca agcgcgaact 60

gtcgcgtgtt accaaggaac gagatttttt acgagacgcg gcagcgtact tcgcgaagca 120

atcgccgaac ggtacgcggt gatcgagcgc tgccgcagcg actaccccat tgggatgatg 180

tgccgctgcc ttcaagtgtc gaccagtggg ttctacgcct gggccaggcg aaagccgggg 240

ccgcgtgccc aggcgaattc gcgtctcttg gagcgcatgc gtgaaatcca cgaggacagc 300

cgaggcatca tcggcgcgcg tcggatgcag gaagacctcg ccgacgaagg catgcccgcc 360

agcttgaatc gggtggcccg cgtcatggcc aaggccgggc ttcagggctg gccgcggcga 420

aagaagcgtg gctttccgcg caagccgccg acgcgtcgtc ccgagggcgt caggaacctt 480

ctggagaggg atttctcggc gctcgaaccg gagacgaagt gggtaaccga catcaccgag 540

atcgtcaccg acgagggaaa actccatctc tgcgtcgtcc tcgacctgta cagcaaactc 600

atcatgggat ggtcgatgca tcaccggcag gatcgccaca tggtggttcg cgcggtacag 660

atggcggttt ggcagcgcga gggcggcgac gaggtgatcc tgcattccga tcgcggcggg 720

cagttcatca gcgatacgta ccagaagttc ctcggcagcc atgccttggt ctgcagcatg 780

agcgaggtcg gccat 795

<210> 15

<211> 156

<212> DNA

<213> virus

<400> 15
ccggcaggat cgccacatgg tggttcgcgc ggtacagatg gcggtttggc agcgcgaggg 60

cggcgacgag gtgatcctgc attccgatcg cggcgggcag ttcatcagcg atacgtacca 120

gaagttcctc ggcagccatg ccttggtctg cagcat 156

<210> 16

<211> 201

<212> DNA

<213> virus

<400> 16
gggctggccg cggcgaaaga agcgtggctt tccgcgcaag ccgccgacgc gtcgtcccga 60

gggcgtcagg aaccttctgg agagggattt ctcggcgctc gaaccggaga cgaagtgggt 120

aaccgacatc accgagatcg tcaccgacga gggaaaactc catctctgcg tcgtcctcga 180

cctgtacagc aaactcatca t 201

<210> 17

<211> 171

<212> DNA

<213> virus

<400> 17
cgcctgggcc aggcgaaagc cggggccgcg tgcccaggcg aattcgcgtc tcttggagcg 60

catgcgtgaa atccacgagg acagccgagg catcatcggc gcgcgtcgga tgcaggaaga 120

cctcgccgac gaaggcatgc ccgccagctt gaatcgggtg gcccgcgtca t 171

<210> 18

<211> 95

<212> PRT

<213> virus

<400> 18

Ser Arg Asp Ser Thr Ile Ala Pro Ala Pro Leu Met Lys Arg Ser Leu
1 5 10 15

Gly Ser Ser Ala Asn Cys Arg Val Leu Pro Arg Asn Glu Ile Phe Tyr
20 25 30

Glu Thr Arg Gln Arg Thr Ser Arg Ser Asn Arg Arg Thr Val Arg Gly
35 40 45

Asp Arg Ala Leu Pro Gln Arg Leu Pro His Trp Asp Asp Val Pro Leu
50 55 60

Pro Ser Ser Val Asp Gln Trp Val Leu Arg Leu Gly Gln Ala Lys Ala
65 70 75 80

Gly Ala Ala Cys Pro Gly Glu Phe Ala Ser Leu Gly Ala His Ala
85 90 95

<210> 19

<211> 47

<212> PRT

<213> virus

<400> 19

Ser Arg Leu Asn His Arg Ser Arg Pro Ala Asp Glu Lys Val Ala Arg
1 5 10 15

Leu Lys Arg Glu Leu Ser Arg Val Thr Lys Glu Arg Asp Phe Leu Arg
20 25 30

Asp Ala Ala Ala Tyr Phe Ala Lys Gln Ser Pro Asn Gly Thr Arg
35 40 45

<210> 20

<211> 275

<212> PRT

<213> virus

<400> 20

```
Met Met Cys Arg Cys Leu Gln Val Ser Thr Ser Gly Phe Tyr Ala Trp
1               5                  10                  15

Ala Arg Arg Lys Pro Gly Pro Arg Ala Gln Ala Asn Ser Arg Leu Leu
            20                  25                  30

Glu Arg Met Arg Glu Ile His Glu Asp Ser Arg Gly Ile Ile Gly Ala
        35                  40                  45

Arg Arg Met Gln Glu Asp Leu Ala Asp Glu Gly Met Pro Ala Ser Leu
    50                  55                  60

Asn Arg Val Ala Arg Val Met Ala Lys Ala Gly Leu Gln Gly Trp Pro
65                  70                  75                  80

Arg Arg Lys Lys Arg Gly Phe Pro Arg Lys Pro Pro Thr Arg Arg Pro
                85                  90                  95

Glu Gly Val Arg Asn Leu Leu Glu Arg Asp Phe Ser Ala Leu Glu Pro
            100                 105                 110

Glu Thr Lys Trp Val Thr Asp Ile Thr Glu Ile Val Thr Asp Glu Gly
        115                 120                 125

Lys Leu His Leu Cys Val Val Leu Asp Leu Tyr Ser Lys Leu Ile Met
    130                 135                 140

Gly Trp Ser Met His His Arg Gln Asp Arg His Met Val Val Arg Ala
145                 150                 155                 160

Val Gln Met Ala Val Trp Gln Arg Glu Gly Gly Asp Glu Val Ile Leu
                165                 170                 175

His Ser Asp Arg Gly Gly Gln Phe Ile Ser Asp Thr Tyr Gln Lys Phe
            180                 185                 190

Leu Gly Ser His Ala Leu Val Cys Ser Met Ser Glu Val Gly His Cys
        195                 200                 205
```

Gly Asp Asn Ala Ala Cys Glu Gly Phe Phe Gly Leu Leu Lys Arg Glu
210 215 220

Trp Ile Tyr Gln Thr Arg Tyr Ser Thr Arg Arg Glu Ala Arg Ala Asp
225 230 235 240

Val Phe Ala Tyr Leu Glu Arg Phe His Asp Pro Arg Met Arg Arg Arg
245 250 255

Val Ala Arg Arg Asp Arg Glu Phe Gln Ala Leu Ile Lys Pro Ser Ala
260 265 270

Glu Thr Gly
275

<210> 21

<211> 69

<212> PRT

<213> virus

<400> 21

Met Val Asp Ala Ser Pro Ala Gly Ser Pro His Gly Gly Ser Arg Gly
1 5 10 15

Thr Asp Gly Gly Leu Ala Ala Arg Gly Arg Arg Arg Gly Asp Pro Ala
20 25 30

Phe Arg Ser Arg Arg Ala Val His Gln Arg Tyr Val Pro Glu Val Pro
35 40 45

Arg Gln Pro Cys Leu Gly Leu Gln His Glu Arg Gly Arg Pro Leu Arg
50 55 60

Arg Gln Arg Ser Met
65

<210> 22

<211> 29

<212> PRT

<213> virus

<400> 22

Met Pro Arg His Glu Val Ala Gly Asn Ile Val Ser Cys Thr Ser Asn
1 5 10 15

Arg Ile Met Ser Ser Arg Ser Phe His Glu Arg Gly Gly
20 25

<210> 23

<211> 29

<212> PRT

<213> virus

<400> 23

Arg Cys Glu Trp Leu Thr Ser Pro Gly Arg Ile Gly Leu Cys Ser Thr
1 5 10 15

Ala Pro Leu Met Thr Asp His Val Leu Leu Arg Ile Met
20 25

<210> 24

<211> 66

<212> PRT

<213> virus

<400> 24

Thr Ser Asn Arg Val Arg Cys Arg Phe Ser Leu Pro Ser Thr Lys Ser
1 5 10 15

Thr Thr Arg Arg Arg Arg Pro Ser Val Leu Phe Thr His Gly Gly Gln
20 25 30

Arg Gly Phe Ala Val Pro Tyr Pro Arg Gly Lys Leu Ala Asn Gly Tyr
35 40 45

Arg Val Pro Val Gly Ser Ala Ser Ala Arg Leu Leu Leu Tyr Cys Arg
50 55 60

Thr Met
65

<210> 25

<211> 37

<212> PRT

<213> virus

<400> 25

His Arg Arg Arg Arg Pro Leu Val Val Asp Glu Arg His Tyr Leu Phe
1 5 10 15

Arg Thr Asp Glu Lys Gly Ala Leu Leu Trp Gln Ile Leu Asp Val Lys
20 25 30

Leu Arg Met Gly Met
35

<210> 26

<211> 28

<212> PRT

<213> virus

<400> 26

Arg Tyr Thr Gly Ser Thr Gly Arg Cys Gly His Arg Pro Arg Cys Cys
1 5 10 15

Ser Arg Pro Arg Gly Asn Arg Arg Cys Arg Leu Met
20 25

<210> 27

<211> 265

<212> PRT

<213> virus

<400> 27

```
Leu Ser Leu Trp Arg Glu Arg Gly Ala Ser Ser Phe Thr Ala Arg Ser
1               5                  10                  15

Leu Arg Ser Ser Asp Arg Thr Val Leu Ser Arg Ser Lys Lys Arg Ser
            20                  25                  30

Ala Ala Ala Tyr Lys Ala Phe Cys Asp Gly Phe Pro Val Arg His Asp
        35                  40                  45

Leu Ala Ala Ala Gln Arg Leu Pro His Trp Asp Asp Val Pro Leu Pro
    50                  55                  60

Ser Ser Val Asp Gln Trp Val Leu Arg Leu Gly Gln Ala Lys Ala Pro
65                  70                  75                  80

Gly Arg Ala Trp Ala Phe Glu Arg Arg Lys Ser Arg Met Arg Ser Ile
                85                  90                  95

Trp Ser Ser Leu Arg Pro Met Met Pro Ala Arg Arg Met Gln Glu Asp
            100                 105                 110

Leu Ala Asp Glu Gly Met Pro Ala Ser Leu Asn Arg Val Ala Arg Val
        115                 120                 125

Met Ala Lys Ala Gly Leu Gln Gly Trp Pro Arg Arg Lys Lys Arg Gly
    130                 135                 140

Phe Pro Arg Lys Pro Pro Thr Arg Arg Pro Glu Gly Val Arg Asn Leu
145                 150                 155                 160

Leu Glu Arg Asp Phe Ser Ala Leu Glu Pro Glu Thr Lys Trp Val Thr
                165                 170                 175

Asp Ile Thr Glu Ile Val Thr Asp Glu Gly Lys Leu His Leu Cys Val
            180                 185                 190

Val Leu Asp Leu Tyr Ser Lys Leu Ile Met Gly Trp Ser Met His His
        195                 200                 205

Arg Gln Asp Arg Pro His Gly Gly Ser Arg Gly Thr Asp Gly Gly Leu
    210                 215                 220

Ala Ala Arg Gly Arg Arg Arg Gly Asp Pro Ala Phe Arg Ser Arg Arg
225                 230                 235                 240
```

Ala Val His Gln Arg Tyr Val Pro Glu Val Pro Arg Gln Pro Cys Leu
245 250 255

Gly Leu Gln His Glu Arg Gly Arg Pro
260 265

<210> 28

<211> 52

<212> PRT

<213> virus

<400> 28

Arg Cys Ser Arg Trp Met Thr Thr Arg Ala Thr Cys Ile Ala Thr Gln
1 5 10 15

Cys Arg Ser Pro Pro Ser Ser Thr Ile Arg Cys Glu Ser Arg Pro Pro
20 25 30

Cys Asn Met Leu Ser Val Tyr Trp Phe Asn Arg Pro Leu Trp Ala Lys
35 40 45

Thr Gln Leu Met
50

<210> 29

<211> 67

<212> PRT

<213> virus

<400> 29

Pro Gln Gly Arg Arg Phe Phe Arg Pro Lys Gly Arg Leu Gly Gly Val
1 5 10 15

Arg Arg Gly Ser Pro Thr Leu Phe Arg Arg Ser Leu Ser Lys Glu Ala
20 25 30

Ser Ser Gly Ser Val Phe His Thr Val Ser Met Val Ser Ile Thr Val
35 40 45

Ser Ser Pro Phe Ser Trp Arg Gln Thr Thr Arg Ser Arg Tyr Leu Leu
50 55 60

Ser Met Met
65

<210> 30

<211> 57

<212> PRT

<213> virus

<400> 30

Ala Gln Ala Leu Arg Phe Gly Pro Gly Arg Ala Trp Ala Phe Glu Arg
1 5 10 15

Arg Lys Ser Arg Met Arg Ser Ile Trp Ser Ser Leu Arg Pro Met Met
20 25 30

Pro Ala Arg Arg Ile Cys Ser Ser Arg Ala Ser Ser Pro Met Gly Ala
35 40 45

Leu Lys Phe Arg Thr Ala Arg Thr Met
50 55

<210> 31

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> amorce sens

<400> 31
cccgccccgc tgatgaaaag 20

<210> 32

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> amorce anti-sens

<400> 32
gcgatggttg agtctcgact a 21

<210> 33

<211> 17

<212> DNA

<213> Artificial sequence

<220>

<223> amorce anti-sens

<400> 33
aggtagcagg cgatatc 17

<210> 34

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> amorce sens

<400> 34
ccttctggag agggatttc 19

<210> 35

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> amorce anti-sens

<400> 35
tgttacctgc tacttcgtgc 20

<210> 36

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> amorce sens

<400> 36
tagagttgcg aggcgtgacc 20

<210> 37

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> amorce anti-sens

<400> 37
ccttatccag tggcttttgg c 21

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.